Modular Design of Biological Systems

by

Julie Erin Norville

B.S., California Institute of Technology (2002)

M.S., Massachusetts Institute of Technology (2004)

Submitted to the Department of
Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Electrical Engineering and Computer Science
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
February 2012

Author .....
Department of
Electrical Engineering and Computer Science
January 9, 2012

Certified by .....
Thomas F. Knight, Jr.
Senior Research Scientist
Thesis Supervisor

Certified by .....
Angela M. Belcher
W.M. Keck Professor of Energy
Thesis Supervisor

Certified .....
Gerald J. Sussman
Panasonic (Matsushita) Professor of Electrical Engineering
Thesis Supervisor

Accepted by .....
Professor Leslie A. Kolodziejski
Chairman, Department Committee on Graduate Students

Modular Design of Biological Systems

by

Julie Erin Norville

B.S., California Institute of Technology (2002)

M.S., Massachusetts Institute of Technology (2004)

Submitted to the Department of
Electrical Engineering and Computer Science
on January 9, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electrical Engineering and Computer Science

Abstract

The focus of my research is the development of technology for building compound biological systems from simpler pieces.

I designed BioScaffold parts, a family of variable regions that can be inserted into a DNA sequence so that at a later time another set of pieces can be substituted for each variable. The variable regions are selective so that a particular piece can be targeted to each region. I have used this technique to assemble protein domains, tune the expression levels of proteins and remove BioBrick scars. BioScaffold parts can be used in combination with BioBrick Standard Biological Parts to create and store devices with tunable components.

I developed simplified methods to produce and examine SbpA, a protein that can either self associate into two-dimensional crystals or bring together fused enzymes when divalent cations such as calcium are added to the protein monomers. My fast and easy purification protocol allows SbpA to be produced under non-denaturing conditions as well as examination of the native state of the protein monomers before crystallization. The absence of a white precipitate when calcium is added to SbpA monomers concentrated to 1 mg/ml provides a simple visual screen that indicates that the protein has failed to crystallize. I also developed a protocol to embed SbpA crystallized on lipid monolayers in trehalose for electron microscopy, allowing creation of a 7 Å resolution map for SbpA.

I created a cells-on-paper system to compose, isolate, and subsequently destack and examine different cell types grown in sheets of ordinary filter paper and maintained in a humidified incubation chamber. I found that *E coli* diluted in LB broth and then applied to filter paper grew at rates similar to the same culture spotted on agar plates. Track etch membranes could be used to isolate different cell types, while still allowing chemical communication between the layers. Use of plasmids that contain fluorescent proteins allowed the behaviour of cells to be tracked using a scanner

after destacking of the layers. The cells-on-paper system can be used both to test and construct modular synthetic systems composed of bacterial ensembles and to create and examine the behavior of compositions of cell types typically found in biofilms.

Thesis Supervisor: Thomas F. Knight, Jr.
Title: Senior Research Scientist

Thesis Supervisor: Angela M. Belcher
Title: W.M. Keck Professor of Energy

Thesis Supervisor: Gerald J. Sussman
Title: Panasonic (Matsushita) Professor of Electrical Engineering

Acknowledgments

I would first like to thank all the institutions who generously provided the grant funding without which this work would not be possible. This material is based upon work supported under a National Science Foundation (NSF) Graduate Research Fellowship, a Vinton Hayes Graduate Research Fellowship, a Massachusetts Institute of Technology (MIT)/Merck Graduate Research Fellowship, a King Abdullah University of Science and Technology (KAUST) Scholar Award, and a grant [355] from the National Academy of Sciences Keck Futures Initiative (NAKFI).

Any opinions, findings, conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

I would like to thank my thesis supervisors for their innumerable contributions to my intellectual development and for inspiration: Angela Belcher for creating a wonderful and happy laboratory that produces beautiful work, Tom Knight for helping found the field of synthetic biology, inventing the BioBrick™ Standard for physical composition (BioBrick assembly), and founding the International Genetically Engineered Machine Competition (iGEM) and the Registry of Standard Biological Parts (Registry), and Gerry Sussman for his mentorship and for reminding me that the biological world is filled with exciting phenomena that one can explore only after the dissertation is complete.

I would also like to thank my collaborators for their contributions to this dissertation, my scientific knowledge, and my intellectual development:

- Chapter 1 and the introductory sections of all subsequent chapters were heavily influenced by conversations with Knight and Sussman.

- Chapter 2 contains the manuscript "Introduction of customized inserts for streamlined assembly and optimization of BioBrick synthetic genetic circuits" [264], written with Ratmir Derda, Saurabh Gupta, Kelly A. Drinkwater, Angela M. Belcher, Andres E. Leschziner, and Thomas F. Knight, Jr.

- Chapter 3 contains the manuscripts "Fast and easy protocol for the purification of recombinant S-layer protein SbpA for synthetic biology applications" [265] and "7 Å Projection map of the S-layer protein SbpA obtained with trehalose embedded monolayer crystals" [263], written with Deborah F. Kelly, Thomas F. Knight, Jr., Angela M. Belcher, and Thomas Walz. Deborah F. Kelly and I made equal contributions to the second manuscript.

- Chapter 4 contains the manuscript "Paper supported polycultures of synthetically programmed bacteria" [266], written with Steven Pennybaker, Ratmir Derda, Matthew R. Lockett, Cory Li, George Whitesides, Angela M. Belcher, and Thomas F. Knight, Jr. In particular I would like to thank Steven Pennybaker for his contributions, which far exceed those of a typical undergraduate researcher, and also Nancy Hopkins, Adam Amsterdam, and Karen Pepper, whose comments substantially shaped the form and structure of that manuscript. I would also like to thank my coauthors Ratmir Derda, Douglas Weibel, and George Whitesides of the NAKFI grant "Using layers of paper to create synthetic microbial communities" [355] that supported this work.

I would also like to thank those who have made my load lighter or my path brighter. Thanks to my parents Scott and Heather Norville and siblings Kristin and Sean for supporting me as a budding synthetic biologist, even before the field was invented, and for supporting my intellectual interest in both biology and engineering.

Thanks to Deborah Hodges-Pabon, whom I met when visiting MIT and who has remained a friend throughout the process. Thanks to Anastacia Feldman for her assistance in the grant process and warm support. Thanks to Norm Margolus, Leslie Kolodziejski, and Donhee Ham for their advice in becoming a researcher. Thanks to Chris and Karen Noren for their advice and mentorship.

I would also like to thank my graduate student friends and comrades Kaity (Katherine) Ryan, Malancha Gupta, Nikku Madhusudhan, Alexey Radul, Max van Kleek, Vivian Lei, and John Hugh Snyder, who also suffered the "slings and arrows of graduate school," and my friends and labmates Deborah Kelly, Danijela Dukovski,

Julius Rabl, Austin Che, Seung-Wuk Lee, Kiley Miller, Nimrod Heldman, Debadyuti Ghosh, Jared Embelton, Jifa Qi, Dong Soo Yun, Mark Allen, Roberto Barbero, Neelkanth Bardhan, Fred Burpo, John Casey, Elizabeth Wood, Sreekar Bhaviripudi, Ahmad Mo Khalil, Eric Krauland, Andrew Magyar, Saeeda Jaffar, and Yu Huang, and all members of the Knight, Belcher, Walz, and Leschziner labs.

Many thanks to the members and advisors of the NSF Synthetic Biology Engineering Research Center SynBERC Student Leadership Committee, including Kevin Solomon, Jonathan Goler, and Gautaum Mukunda.

Thanks, too, to the organizers of the supplemental courses that accelerated my development as a researcher and academic: the Protein Purification and Phage Courses at Cold Spring Harbor Laboratory, the Electron Microscopy Course at the European Molecular Biology Lab, and the Computing Beyond Silicon Summer School developed by André DeHon and Erik Winfree at Caltech that introduced me to synthetic biology and the work of Thomas Knight and Angela Belcher.

Finally, special thanks to my friends, Samuel Freilich, Michelle Kats, and D. J. Gallagher, for their continued friendship and support.

Contents

List of Figures

List of Tables

Chapter 1

Modular Design of Biological Systems

How can modularity be used to engineer organisms? Let us consider a conceptual example. Many insects exist on similar length scales (see Figure 1-1 on the following page). Imagine that we could transfer anatomical components from a beetle to a butterfly or that we had knobs and switches that would allow us to control the size, shape, or function of an insect's anatomical elements. These modular transformations, as fantastic as they seem, describe the type of changes that it may be possible to incorporate into natural biological systems using genetic engineering. The organization of natural biological systems in some cases facilitates the emergence of these types of changes [21, 292, 348]. By harnessing the ability to make desirable changes on a fast time scale, we can gain the ability to engineer biology in a systematic and straightforward fashion.

This dissertation demonstrates how modularity can be introduced into the construction of biological systems in three different domains: into linear DNA sequences, into two-dimensional (2D) protein structures, and into three-dimensional (3D) multicellular ensembles. After further examination of modularity in Chapter 1, Chapter 2 on page 55 describes a component for modular construction of DNA sequences. BioScaffold parts (or BioScaffolds) are a family of components that can be introduced into genetic sequences and then replaced at future points in time. BioScaffolds are

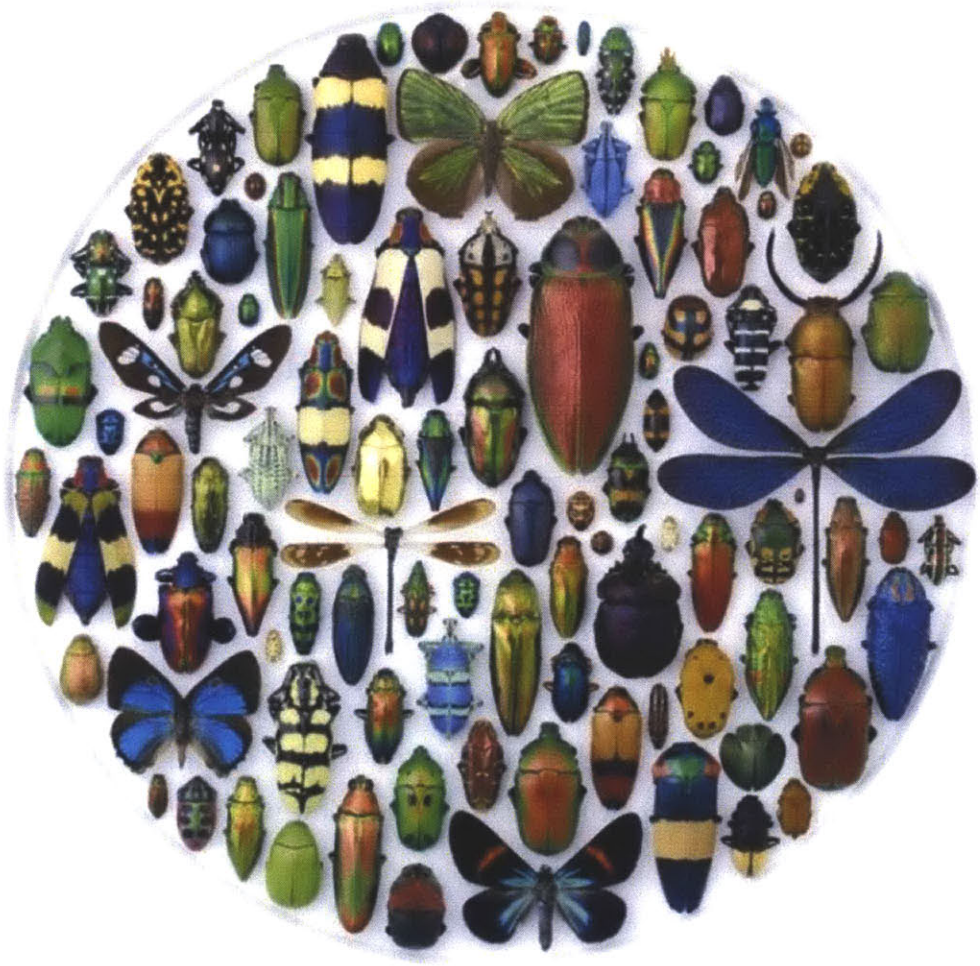


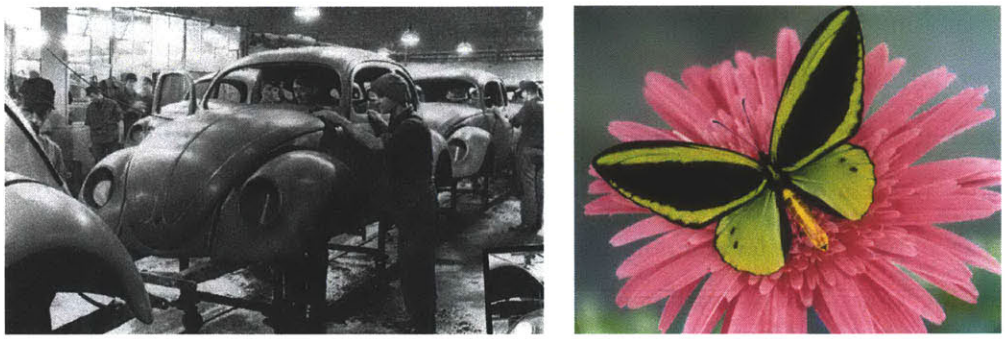
Figure 1-1: Natural Examples of Modular Design: Insects. The artwork Aesthetica Sphere by Christopher Marley [73] displays a variety of similarly sized preserved insects. Imagine that it would be possible to exchange the wings of different insects or change the size or function of individual features.

powerful because they allow future modification of both themselves and surrounding DNA sequences. Chapter 3 on page 79 explores at a deep structural level a modular component for construction of protein structures as well as a simple way to produce the component. The surface-layer (S-layer) protein SbpA is remarkable because it can be controllably switched from a soluble, flexible, and monomeric state

to a self-assembling and then a rigid and supramolecular state (where it forms 2D protein crystals), upon addition of calcium or other divalent cations. Chapter 4 on page 115 introduces a technology for creating 3D and structured multicellular ensembles. Multiple bacterial strains can be assembled into a structure, by inoculating pieces of filter paper and then stacking them together. Interleaved membranes isolate different strains, but allow molecular communication between the layers. Unstacking the paper sheets at a later time is a simple process that allows one to assess cellular behavior throughout the structure.

1.1 Introducing Modularity into Engineered and Natural Biological Systems

Modularity is a fundamental design principle. Modularity is present throughout the engineering and natural biological world [181, 306]. An engineered example of a modular system is an assembly line for car manufacture (Figure 1-2(a) on the following page), in which each worker has a different specified function. A natural example of a modular system is a butterfly (Figure 1-2(b) on the next page), in which different anatomical components have distinct functions. Modularity is the fundamental way in which systems are divided into parts. Modularity allows you define carefully what the interfaces of a component are. Modular systems are less complex because they are subdivided into functional components that perform specific functions [41]. The division of the elements of a design into smaller functional components decreases complexity [41]. Modular systems are flexible. Modular systems, built from mix-and-match part families, can be adapted to new problems quickly and easily. In the sections that follow, I will describe several examples of modularity.



(a) Modular Assembly Line. (b) Modular Anatomy of an Insect.

Figure 1-2: Examples of Modularity. Modularity is present in both engineered (a) [19] and natural biological systems (b) [11].

1.2 An Example of Engineered Modularity: Containerization for Global Freight Transport

Before we begin to examine modular biological systems, we will first examine the shipping container, an example of how modularity is used in engineering. Malcolm McLean invented the system that became International Organization for Standardization (ISO) certified standard containers in 1956 [242, 285], streamlining the process of freight transport. Containerzation allows compartments filled with any desired material to be brought together and arranged using a standard interface. In this dissertation, we will introduce the cells-on-paper platform, which analogously allows us to bring together cellular compartments that interact along standard interfaces.

1.2.1 Shipping Containers before Containerization

Bulk-break shipping, a far less efficient system, preceded containerization. In bulk-break shipping, goods were placed in drums, crates, or boxes of non-uniform size [102] or were lashed to skids or pallets [5]. In order to load or unload goods, the barrels or crates were tied together with rope and transferred by hooking the rope onto a crane. After unloading, the goods were transferred using handcarts [27]. "[C]argo was manipulated, piece by piece, by muscle, a longshoreman's hook, rope slings, four wheelers, the lever, the pulley, the inclined plane, or any means possible" [167]. Over-

all, the process was so labor intensive and time consuming that ships spent at least "half of their time in port," even after the introduction of the forklift in 1935 [339]. The large number of workers, the identifiable contents of non-uniform containers, and the many small packages in bulk-break shipping, made goods subject to theft. A certain percentage of goods were often "set aside" for the dockworkers' supervisors [244]. Also, since the packages used in bulk-break shipping were not weatherproof and were subject to theft, they needed to be enclosed within vehicles during transport and in warehouses during storage [308]. In comparison, containerization utilizes and demonstrates the benefits of modularity at multiple levels of hierarchy by decreasing the costs of packaging, crating, loading, unloading, and loss or damage of cargo [308].

1.2.2 Modularity Introduced by Containerization

Standard shipping containers are modules used in international freight transport. Different types of shipping containers can have diverse function. The length, strength, fittings, and internal and external surfaces [214] create the interfaces between the contents of a shipping container and the external environment that surrounds it. The twistlock fittings on shipping containers allow them to be connected to other structures (such as cranes, vehicles, and other containers) that enable the container to be moved over land and sea as well as stacked in a variety of configurations [62]. Although container appearance is fairly uniform, permanent identification marks called ISO codes on the exterior (that both identify the container owner and uniquely label the box) allow containers to be tracked and routed to their final destination without being opened or inspected at every port [157, 253]. The seals on closed containers are also labeled with a code, but in this case the identification mark is only valid for a single shipment [157]. Computer-based systems are now used to track individual containers in real time using GPS [244].

Different Types of Standard Shipping Containers and Their Functions

Modular systems, such as standard shipping containers, are flexible. Different types of standard shipping containers are optimized to transport goods that have different shapes, sizes, and properties. Since freight is enclosed in standard shipping containers, transport vehicles can transport many different kinds of goods at the same time. Many different types of shipping containers exist, though each container has a standard length and fittings. General-purpose (or dry cargo) containers are typically used to transport boxes, cartons, cases, sacks, bales, pallets, and drums [13, 157]. Although the exterior form factor and fittings of containers must remain fixed, the internal components of containers can be altered to tailor them for different purposes. First, some types are optimized to facilitate different types of loading. Open end containers allow loading and unloading of standard size pallets [13]. Open side containers allow loading of oversized pallets [13]. Open top containers allow direct loading of materials such as logs, grain, and ore [13]. Second, some types are optimized to transport specific classes of goods. Tanktainers (or tank containers) are designed to hold liquids or gases and contain a tank and a framework that surrounds it [13, 88, 157, 250]. Reefers (or refrigerated containers) maintain the quality of perishable goods by controlling the storage temperature [13, 18, 157]. Garmentainers (or garment containers) store clothing on hangers, allowing it to be transported in an unwrinkled state [14]. Third, containers can be optimized for reuse. The height of collapsible containers can be reduced by 3/4 when they are empty, reducing the cost of transporting them to locations where they are needed [313].

Interfaces of Standard Shipping Containers

The interfaces of standard shipping containers have been carefully designed. Steel shipping containers must meet ISO specified strength requirements (e.g., for stacking load, lifting force, restraint force, and racking force) [88]. In addition, closed containers (such as the dry cargo containers are designed to be weatherproof [88]. The handling of a container and how it is stacked and unstacked either for storage or onto

a transport vehicle is identical for each container. The modularity of the container breaks up the cargo into units of fixed size and allows the containers to interact and be combined into larger units [214]. The fittings present on containers are holes called corner castings. A device called a twistlock is placed in into each of the upper corner castings of a container, before a new container is stacked on top. After placement of the new container, the twistlocks on each corner are turned, locking the two containers together. This process can be repeated until the stack of containers is as high as desired. In addition to being stacked on top of other containers, standard shipping containers can also be linked to a transport chassis (e.g., a container ship, a truck, or a rail car) [244]. On the decks of ships, containers can be stacked seven high above deck and eight deep below deck [325]. On rail cars or trucks, the stacks are sometimes one and often two containers high [214]. When stored at ports, containers are often stacked on the ground in blocks that are up to four containers high or may be placed individually on rubber-tired trailer chassis [171]. The strong, weatherproof exterior of the dry cargo container protects the enclosed goods from damage and eliminates the need for a warehouse. The corner castings on shipping containers are also used to attach the containers to cranes or other handling vehicles that move containers between storage and the transport chassis [244]. Most cranes and handing vehicles contain a special device called a spreader that attaches to the four corner castings at the top of the container [290]. Some handling vehicles without spreaders, such as sidelifters, use chains attached to the four corner castings at the bottom of the container [10]. If a container is placed on a trailer, non-specialized vehicles can be used to pull the cargo. Different handling cranes and vehicles are used for different purposes, whether to stack containers to particular heights or onto a different transport chassis.

Transport Routing of Standard Shipping Containers

Modularity decreases complexity by dividing systems into parts. Shipping containers simplify the transport routing of goods, making it easier to both efficiently sell space on cargo vehicles and sort containers at intermodal port, rail, or road terminals [66, 325]. In containerization, standard shipping containers of fixed size contain a

large number of smaller goods that have the same final destination. Dissemination, sorting, and routing thus can be performed for containers rather than individual goods [325]. As mentioned previously, standard shipping containers, have similar exteriors, but are distinguished by an identification mark. Unfortunately, however, when identification marks are copied manually, 10% to 20 % of the recorded codes may contain errors [165]. Automated methods of tracking containers, such as computer recognition of the marks painted on containers; RFID tags on goods, containers, and tracking vehicles; and GPS tags both simplify audits of goods at each location and help identify the locations where transport routing errors and thefts occur [165, 244, 247]. The similarity between standard shipping containers originally made identification of valuable theft targets more difficult, but currently there is hope that smart tags can aid in pinpointing the precise locations where thefts, losses, and routing errors occur [247]. When a single container full of goods is lost or stolen, (rather than the constant stream of diverted merchandise during the age bulk break shipping) the costs can still be high [244]. Since the machinery for freight transport and unloading only requires a few operators, opportunities for theft by a large number of individuals are reduced.

1.3 An Example of Natural Modularity: Compartmentalization in Biological Systems

Now we will explore how modularity is used in biological systems. Although some types of modularity are present in prokaryotic systems, modular design is used extensively and at multiple levels of hierarchy in eukaryotic systems. The fundamental modular unit in eukaryotic systems is the membrane-enclosed compartment. The use of membrane-enclosed compartments as modules allows both unicellular and multicellular eukaryotes to overcome diffusion barriers, providing spatial and temporal control of the distribution of molecules [351]. Multicellular eukaryotes are larger in size than unicellular eukaryotes, which are bigger than prokaryotes. The size differences are due to the ability to overcome diffusion barriers at different length scales. Unicellular and

multicellular prokaryotes are limited in size because they do not activity control the diffusion process within the cell [245], although they can localize proteins to specific positions [122, 235, 310] and control diffusion across membranes [199]. A few types of giant bacteria exist [245, 307]. The largest recorded example is *Thiomargarita namibiensis* (*T. namibiensis*), a coccoid bacteria that can reach 0.8 mm in diameter, (a size that is visible to the naked eye and over a million times larger than *E. coli*) [209]. They typically have multiple nucleoids, or bacterial chromosomes, dispersed throughout their interiors and/or large internal vesicles that store food and reduce the area of the cytoplasm [245]. Multicellular stromatolites or bacterial mats typically contain multiple types of bacteria that exist in consortia where members each create nutrients for the other species to consume [272]. Within these films, diffusion is passive. Unicellular eukaryotes control diffusion using vesicular transport [324]. Gigantic single-celled eukaryotes, such *Syringammina fragilissima* (*S. fragilissima*) can reach 25 cm in diameter, have multiple nuclei and in some cases consume prokaryotes that grow on their surfaces [349]. Multicellular eukaryotes contain cell types, tissues, or organs that aid in the acquisition, transport, and removal of chemicals. For instance, in mammalian systems the respiratory and circulatory system obtains oxygen and removes carbon dioxide whereas the digestive and excretory system aids in the acquisition of nutrients and the removal of metabolic waste products [117].

1.3.1 Compartmentalization in Single-Celled Prokaryotes

Single-celled prokaryotic and eukaryotic organisms must perform the same functions. To remain alive these organisms must adapt, retain their structure, perform metabolic processes, maintain homeostasis, and reproduce [127]. Eukaryotic and prokaryotic cells each contain compartments that aid them in performing these functions [176]. However, prokaryotic compartments are diverse in their form and structure, while eukaryotic compartments are modular and membrane-enclosed. In this section we will explore the different types of prokaryotic compartments. One of the compartments that we describe are the bacterial organelles known as microcompartments. Bacterial microcompartments are protein cages that enclose and bring together enzymes for

metabolic pathways. In this dissertation we will create techniques for the S-layer protein SbpA, a protein which naturally creates protein cages around the outside of cells, that will simplify its use in bringing together proteins in close proximity. Crystallization of SbpA allows it to bring together fused enzymes using a standard interface.

Physical Interfaces within Prokaryotes

Prokaryotes contain diverse physical enclosures for compartmentalization [221], though these enclosures are quite different from membrane-enclosed enclosures in eukaryotes. Some bacteria have genetic programs for protein-enclosed microcompartments. Microcompartments can be used to sequester metabolic reactions [363]. For example, carboxysomes concentrate carbon dioxide for the enzyme Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) [183]. Gas vesicles are another type of protein-enclosed compartment that increases the buoyancy of cells [350]. Prokaryotes can also sometimes have fully or partially membrane-enclosed compartments [107, 197, 312] such as volutin granules, chlorosomes, and magnetosomes. Prokaryotes also exhibit external compartments [221]. In most bacteria, a rigid layer of polysaccharide such as peptidoglycan surrounding the plasma membrane determines the shape of the cell [248]. (The structural role of peptidoglycan was proven by experiments, which showed that bacteria form round spheroplasts after the removal of the peptidoglycan-layer but that the peptidoglycan-layer maintained its shape even after removal of the contents of the bacterial cell [156].) Gram-positive bacteria have an exterior peptidoglycan layer [360]. Gram-negative bacteria are surrounded by peptidoglycan, a periplasmic space, and an outer membrane [360]. Gram-positive bacteria, gram-negative bacteria, and even archaea can be surrounded by an exterior, crystalline protein coat called the S-layer [303], though this structure often is absent when the organisms are cultivated under laboratory conditions [319]. Although the purposes of the S-layer have not yet been completely elucidated, they are believed to help microorganisms adapt to different environmental conditions [318]. Some microorganisms change the protein used to create the S-layer due to exposure to different temperatures and gas

concentrations [303]. In some cases, metabolic enzymes such as amylase are linked to the S-layer [112]. The S-layer may also be used to evade external threats, such as immune responses and phage [319], since they also control interactions with these external variables.

Spatial Interfaces within Prokaryotes

Prokaryotes also exhibit many types of spatial compartments where segregation of specific molecules occurs in the absence of physical compartments. The nuclear material in bacteria, for instance, is segregated by molecular crowding, DNA supercoiling, and binding of architectural proteins to a central region of the cell in a region called the nucleoid [234], even though the space is not enclosed by a membrane-enclosed compartment. In some bacterial cells, transcription is localized to the nucleoid and translation to the cytoplasm [218], which mirrors the separation of these processes in the nucleus and cytoplasm of eukaryotic cells [280]. A more universal feature is that the nucleoid is highly structured and domains occupy specific regions of the cell at set points in time relative to the cell cycle [260]. A recent study has revealed that mRNA transcribed from the nucleoid is localized at specific locations within the cell, even though models predict free diffusion [226]. This characteristic means that protein synthesis is also localized to specific locations, perhaps aiding in complex formation and contributing to the colocalization of proteins created by translation of genes on an operon. Elowitz has shown that large complexes can become localized within bacterial cells [115]. Thus, spatial partitioning of large protein complexes may occur in prokaryotic cells, even in the absence of membrane-enclosed compartments.

Molecular Transport Routing within Prokaryotes

Transport of molecules by prokaryotes is both an active and a passive process that occurs both within and across compartments. Most movement within compartments occurs through diffusion [299], though active transport of specific molecules across compartments is common [199]. In comparison, eukaryotes have mechanisms that allow active transport within the cell [324]. Alternatively, proteins that can diffuse

freely may later become restricted to certain locations due to affinity capture [299]. Some features, such as microcompartments [304], plasmids [31], and magnetosomes [197] are attached to a simple actin-like bacterial cytoskeleton. Signal tags can have affinity to the membrane or to molecular machinery that either covalently attaches the protein to the membrane or translocates it across the membrane. In addition, some types of molecular machinery can even secrete the protein into another organism [241]. Within the cell, a signal tag that facilitates entry into a bacterial microcompartment has been discovered [271]. Bacterial feeding is also mediated by active transport and diffusion of molecules into the cytoplasm of cells. On the outside of bacterial cells, free or cell-associated enzymes break down macromolecules [50, 206]. As described previously, molecules that diffuse to the plasma membrane are transported across it using membrane proteins.

1.3.2 Compartmentalization in Prokaryotic Cellular Aggregates and Consortia

Biological cells are the simplest unit of self-contained biological life, and themselves can act as modular subunits within larger ensembles. In prokaryotes, biological cells can be maintained individually in a planktonic form but they can also grow together in a multicellular composite called a biofilm [328]. Within a biofilm, prokaryotic cells display some characteristics of multicellularity [309].

Cellular ensembles of both prokaryotes and eukaryotes must solve many of the same challenges. They must create organized structures that consist of multiple cells, maintain structural stability and form beneficial diffusion gradients of gases and molecules across the structure. The formation of consortia yields a number of advantages over single-celled forms for prokaryotes. Multicellular ensembles can have a division of labor across multiple cell types, access new environmental niches, create collective defenses against threats, and develop new techniques to optimize population survival [309]. Prokaryotes can exist as clumps, filaments, hollow spheres, or sheets of varying thickness (such as biofilms, microbial mats, and stromatolites)

[54, 184, 186]. In modern times, the majority of macroscopic prokaryotic consortia are small, since metazoans tend to consume them or poison them by introducing oxygen into obligate anaerobic layers before they achieve larger forms [323]. Large microbial mats and stromatolites, however, still exist in extreme conditions such as hypersaline, anoxic, or high-temperature aqueous environments [269]. Before the evolution of tunneling metazoans, the fossil record indicates that stromatolites were the dominant macroscopic life form on earth [323].

Physical Interfaces between Prokaryotes

Prokaryotes can form aggregates through either direct physical connections between cells or as a collection of multiple individual cells in consortia; however, these aggregates do not take on the properties of true multicellular organisms like the eukaryotes described in Section 1.3.4 on page 38. We find that true multicellularity only occurs in eukaryotic systems and in cases where mechanisms that allow active transport of nutrients are present [75].

Either physically connected cells or the individual cells within consortia can be surrounded by an extracellular polymeric substrate (EPS). EPSs can be composed of proteins, DNA, polymers, sediments, or precipitated inorganic minerals [110, 126] and provides additional protection against desiccation, predation, and other threats [361]. One example is that filamentous forms of physically connected bacteria that are surrounded by EPS sheaths [329]. A second example, stromatolites are lithified consortia of prokaryotes encapsulated in calcium carbonate or other minerals [110]. Some sheetlike consortia that include aerobic or facultative anaerobic bacteria contain internal water channels facilitate to gas and nutrient exchange [327]. In comparison, I described in the prior section how channels created by tunneling metazoans could destroy stromatolites by poisoning anaerobic layers with oxygen. Thus, stromatolites tend to contain primarily vertically stratified layers.

Prokaryotes can also form aggregates when they septate in one, two, three, or many random planes but do not divide [322]. Typically a species with this property characteristically forms one of the following types of clusters: two-cells (diploids),

chains or filaments (streptoids), four-cells (tetrads), a cube of eight-cells (sarcinae). and grape-shaped clumps (stayphyloids) [58]. For example, *Deinococcus radiodurans* (*D. radiodurans*) is a species whose cells form tetrads through "alternate and synchronized divisions in two planes" [255]. Another example of a complex connected arrangement is the hollow spheres formed by obligate multicellular phototactic or magnetotatic prokaryotes [186, 210]. These obligate multicellular prokaryotes reproduce by division into two new hollow spheres. As the processes of septation and cell division are under genetic control, physically connected aggregates can be derived from mutations in single-celled prokaryotes. For instance, the envA mutant of *E. coli* forms filaments by failing to divide [261].

In some cases, individual cells in a physically connected structure may exhibit heterogeneous functions, in a manner reminiscent of the differentiated cells in multicellular eukaryotic organisms. For example, one cell or a small number of cells may anchor a filament to a surface, especially in high flow conditions (i.e., this occurs in mice and rats where a holdfast segment anchors a chain of segmented filamentous bacteria to intestinal ilea mucosa cells [179].) A second example is a subpopulation of cells that produces a desired metabolic product for surrounding cells (i.e., when filamentous cyanobacteria are deprived of nitrogen, regularly spaced cells along the chain become heterocyst cells that fix nitrogen for their neighbors [23, 358].)

Spatial Interfaces between Prokaryotes

Sheetlike consortia of prokaryotes often contain stratified layers that contain organisms with different metabolic properties [229]. In the consortia, the top layer is typically exposed to an air or water interface and the bottom layer is adjacent to a physical substrate. Biofilms consist of individual cells or aggregates such as filaments that form a pure or mixed-species culture. In the initial phase of growth, a biofilm, microbial mat, or stromatolite may consist of a layer of cells in a single metabolic state (i.e., the pioneer colony) [202]. As the sheet grows vertically upward from the substrate new microenvironments emerge that are typically populated by bacteria in different metabolic states. In a mature biofilm, a common configuration may be a

three-layered structure that contains a nutrient and oxygen rich top layer, a nutrient rich and oxygen poor middle layer, and a nutrient and oxygen poor bottom layer [328]. An example of multispecies consortia is microbial mats from the hypersaline Ebro Delta of Spain [259]. The top layer contains light, oxygen and cyanobacteria that perform aerobic photosynthesis. The middle layer contains light but no oxygen and is populated by green and purple sulfur bacteria. The bottom layer contains no light or oxygen and is populated by sulfate-reducing bacteria that decay detritus and dead cells that have been left behind during vertical growth. In consortia of single species such as biofilms containing an obligate aerobe such as *Pseudomonas aeruginosa* (*P. aeruginosa*), a mature biofilm also develops chemical gradients: for instance, a top layer that contains nutrients and oxygen, a middle layer that contains nutrients but is oxygen poor, and a bottom layer that is nutrient and oxygen poor. Experiments utilizing *P. aeruginosa* biofilms demonstrated that only the oxygenated layer was metabolically active; however, changing the atmospheric concentration of oxygen altered the thickness of the layer [362].

Molecular Transport Routing across Prokaryotes

Movement of molecules such as chemicals and gases typically occurs by diffusion and is characterized by sharp gradients [133]. In small clumps of physically connected cells or hollow spheres the concentration of nutrients, wastes, and gases may not differ substantially from those in single bacteria cells. For example, extending the length (rather than radius) of a filament of physically connected rod shaped cells does not affect a cell's surface to volume ratio (and thus either diffusion of nutrients of wastes) [196]. Similarly. the obligate multicellular bacteria that form the sphere that divide during reproduction maintain their surface-to-volume ratio within strict boundaries [186]. In consortia such as biofilms, mats, and stromatolites, nutrient cycling occurs when metabolic wastes from one layer of bacteria become nutrients that are consumed by another layer of bacteria [94]. Similarly, physically connected filaments of cyanobacteria contain a certain ratio of heterocysts create nitrogen, which diffuses into neighboring cells, as was described previously. Diffusible signaling molecules,

such as N-acyl homoserine lactone (HSL), can provide bacteria in biofilms with information that is used to trigger group behaviors. For instance, the plant pathogen *Erwinia carotovora* (*E. carotovora*), does not secrete digestive enzymes until they have reached a certain concentration (i.e., the concentration of HSL reached a certain concentration) [293]. Enzymatic digestion of nutrient molecules allows them to diffuse into the biofilm.

Alternately, aggregated prokaryotes or individuals within consortia may themselves exhibit movements, which typically occur to increase their ability to obtain nutrients. Some types of clumps and filaments as well as hollow spheres of phototactic or magnetotactic bacteria exhibit directed group movement [186, 210]. Cells in sheetlike consortia also exhibit characteristic movements. The most common type of movement is vertical accretion of cells through growth, a process that occurs in biofilms, microbial mats, and stromatolites. As the height of the top level of the microbial mat (from the Ebro delta) increases, sulfate-reducing cells that perform decomposition migrate into layers that previously contained living green and purple sulfur bacteria but now contain detritus and dead cells [259]. Some types of filamentous cyanobacteria continually adjust their height in a stromatolite in order to obtain the optimal light intensity [353]. In addition, bacteria or multicellular aggregates can be released from bacterial sheets when they are exposed to high sheer forces, providing a route to seed new colonies [202]. Also, some biofilms can also be programmatically dispersed into planktonic forms. *P. aeruginosa*, which populates the lungs of cystic fibrosis patients, will disperse if the concentrations of HSL reach a certain level [103]; however, the flow of blood in the capillaries that surround the alveoli causes the signal to diffuse away.

1.3.3 Modular Compartmentalization Using Membrane-Enclosed Compartments in Eukaryotic Cells

Membrane-enclosed compartments are modules within eukaryotic organisms. The membranes create the interface between the internal contents of a compartment and

its external environment. Membrane-enclosed compartments within cells move when they are attached to the cytoskeleton using adaptors and molecular motors that direct the movement of vesicles to different locations [298]. Membrane-enclosed compartments within cells can also undergo several different standard transformations that are used in different ways throughout the cell [124]. In the next section (Section 1.3.3), which focuses on modularity in multicellular eukaryotes, I will describe genetically encoded spatial compartments that contain arrangements of membrane-enclosed cells. Membrane-compartments have properties that are analogous to the BioScaffold parts that we will introducc in this dissertation. Membrane-enclosed compartments may contain a variety of substances internally; however, their exterior interface remains the same. BioScaffold parts can be assembled with BioBrick parts to create new BioBrick parts. At a later time, the BioScaffold part can be replaced with a new sequence. The exterior interfaces of the BioBrick part that contains the BioScaffold, however, will remain the same.

Types of Membrane-Enclosed Modules within Eukaryotic Cells

Modular systems, such as the membrane-enclosed compartments in eukaryotic organisms systems, are flexible. First, eukaryotic cells contain multiple types membrane-enclosed compartments [144]. Organelles are large membrane-enclosed compartments. Vesicles for endocytosis, exocytosis, or intracellular transport are small membrane-enclosed compartments. Second, different membrane-enclosed compartments have different functions. For instance, the nucleus contains the chromosome and is the region where transcription and splicing occur, lysosomes digest particles imported into the cell as well as old organelles, and the Golgi-apparatus packages molecules such as proteins or lipids into vesicles for transport within or outside the cell [227]. In addition, pinocytosis creates small liquid-containing vesicles, whereas phagocytosis yields large solid matter containing vesicles [95]. Third, the types and quantities of membrane-enclosed compartments contained within membrane-enclosed cells are diverse. For instance, plant cells, but not animal cells, contain chloroplasts. Alternatively, during erythropoiesis, "the nucleus, ribosomes, lysosomes, endoplasmic

reticulum, and Golgi apparatus" [291] are either destroyed or removed during the maturation of a red blood cell. Although all mechanisms have not been elucidated, the nucleus is extruded and the mitochondria destroyed by autophagy.

Transformations at Interfaces of Membrane-Enclosed Compartments

The modularity of membrane-enclosed compartments decreases complexity in eukaryotic cells. Eukaryotic cells contain molecular machinery that can implement several vesicle transformations that are used in multiple ways throughout the cell [59]. First, as described above, vesicles can become attached to molecular motors that move them along the cytoskeleton to different locations in the cell. Second, vesicles can bud off and form from existing membranes. For example, protein- or lipid-containing vesicles also bud off from the membrane of the Golgi apparatus. Vesicles can bud off and form from the plasma membrane during endocytosis (i.e., pinocytosis and phagocytosis.) Researchers believe that mitochondria and chloroplasts are derived from organisms (i.e., alpha purple bacteria and cyanobacteria) that were absorbed by endocytosis, became endosymbionts, and eventually became controlled by the nucleus [139, 268, 342]. (Researchers speculate that transfer of control of these organelles to the nucleus is believed to have happened due to lysis of a portion of those cells, and the transfer of their DNA to the nucleus [240].) Endocytosis may not have been required for this process, however, as symbiotic internalization of one cell within another has also been observed in prokaryotes [344]. In a similar process, the single-celled and amoeboid eukaryotes called foraminiferans consume other organisms. The foraminifera then cytoplasmically maintain either the cells (typically bacterial or algal) as endosymbionts or the algal chloroplasts (obtained through kleptoplasty) as organelles for long periods of time within the cytoplasm [254]. (Kleptoplasty, with long-term utilization of the ingested chloroplasts, has also been observed in ciliates and Sacoglossan sea slugs [300].) Third, vesicles can fuse to existing membranes. Vesicle fusion allows vesicles that contain proteins for secretion to fuse with the plasma membane and secrete their contents outside the cell [208]. Vesicle fusion allows proteins that would be harmful to the cell to be transported directly to and fuse with the digestive lysosome after

secretion from the Golgi apparatus [301]. Vesicle fusion allows a vesicle that contains an endocytosed liquid or solid to be fused to the lysosome [143]. Vesicle fusion also allows old or damaged organelles to be fused with the lysosome for destruction during the process of autophagy [185].

Transport Routing of Membrane-Enclosed Compartments

Membrane-enclosed compartments have been evolved to implement specific functions. Movement of organelles and vesicles is mediated by molecular motors that cause these compartments to move along the cytoskeleton [80]. Vesicles containing food or liquid from from the plasma membrane to the lysosome. Vesicles containing newly created proteins move from the endoplasmic reticulum to the Golgi apparatus and sometimes back again [59]. The Golgi apparatus serves as a post office of the cell, shipping membrane-enclosed vesicles that contain proteins and lipids to organelles or the exterior of the cell [201]. Some secretory vesicles must receive signals before they will release their contents to the exterior of the cell [143]. As proteins are translocated into the endoplasmic reticulum during translation and then packaged into vesicles by the Golgi apparature, proteins can be moved by active transport of the vesicles to other membrane-enclosed compartments regardless of the size of the proteins or the complexes that they form [64, 217]. This means that in eukaryotes it is possible to easily transport large complexes within or outside the cell. Membrane-enclosed compartments in eukaryotes also contain machinery to translocate signal-containing proteins across their exterior and interior membranes [305]. The membrane-enclosed compartments within a cell can sometimes be derived by the transfer of other organisms to the cytoplasm of a mixotrophic host, as is described in Section 1.3.3 on the preceding page for foraminiferans. (In addition, researchers have speculated that ingested intracellular endosymbionts and chloroplasts may serve as local sinks for waste metabolites, helping foraminiferans achieve large sizes by overcoming transport problems [146].)

1.3.4 Modular Compartmentalization Using Genetically Defined Spatial Containers in Multicellular Eukaryotes

Modular design decreases complexity. Due to the use of modularity in patterning of organisms, only a small number of genes are required to regulate the process. Homeobox genes aid in the structural and temporal development of patterned regions in metazoans, plants, and fungi [134]. Metazoans have the Hox genes, which are transcription factors that either up or down regulate other genes that control development [273]. For example, homeobox genes allow functions that are present in single-celled protozoans to be distributed over multiple cells in metazoans [79]. For example, the paramecium is an entire complex organism present in a single cell. Food first enters the paramecium's gullet from the oral groove; second, the nutrients are encapsulated into food vacuoles; third, the food vacuole is fused with the lysosome; and finally, any remaining matter is excreted through the anal pore [116, 236]. Homeobox genes allow the digestive structures similar to that of the paramecium to be writ large and over multiple cell types in more complex metazoans. The digestive function, can be compartmentalized into a specific cell type, a tissue, an organ, or an organ system, depending upon the complexity of the metazoan [123]. The digestive cells, tissues, organs, or organ systems are each modular in their design [79].

The modularity of membrane enclosed compartments in eukaryotes probably eased the transition to multicellularity. In Section 1.3.2 on page 31, I described how prokaryotes can take on certain properties of multicellular organisms; however, in eukaryotes we can actually observe a transition from a single-celled organism to a multicellular organism. The green-algae *Chlorella vulgaris* (*C. vulgaris*) is typically present in a single-celled form; however predation by the protozoan *Ochromonas vallescia* (*O. vallescia*) causes the cell to transition to an eight-celled form that is retained in all subsequent generations [60]. The globose form prevented predation, yet allows nutrient diffusion from the external environment [60]. This assembly in some ways is similar to the bacterial aggregates described previously; however, the difference between eukaryotes and prokaryotes is that they have natural machinery

that allows them to use membrane vesicles both for internal transport and to interact with the plasma membrane. Therefore, it is easier for eukaryote cells to optimize transport across larger collections of cells and for both predators and prey to increase in size and complexity in order to fill different environmental niches; whereas the structure of bacterial systems is determined by diffusion gradients. Each of the eukaryotic clades that developed multicellularity had machinery for actively transporting nutrients between cells (i.e., plasmodesmata, gap junctions, or incomplete cell walls) [75].

Homeobox Genes Divide Multicellular Eukaryotes into Spatial Containers

By dividing multicellular eukaryotes into spatial containers, homeobox genes create a structure that regulates downstream developmental processes [273]. In metazoans, the number of Hox or Hox-like genes present depends upon the complexity of the organism. Placozoans are simply structured, containing only four cell types and one Hox-like gene [170]. In the pancake-shaped placozoans, the Hox-like gene is expressed at the edge of the organism where cells divide into the three layers that create the upper epidermis, syncytium (or fiber cells), and lower epidermis. In the radially symmetric cnidarians, multiple Hox genes are expressed along the oral-aboral axis of the organism [125]. In arthropods, a linear array of Hox genes is expressed along the anterior-posterior axis of the organism [93]. In most vertebrates, the linear array of Hox genes has been duplicated twice relative to invertebrates (though an additional duplication has occurred in modern-ray finned fish,) and the Hox genes are expressed along the anterior-posterior axis of the organism [246]. Thus, in most metazoans, Hox gene expression occurs along the anterior-posterior axis of the organism [69]; however, tetrapods, the vertebrates that evolved from bony fishes and include mammals, amphibians, reptiles and birds, (and some of their fish ancestors) have an unusual Hox gene expression structure. The majority of Hox genes are expressed along the anterior-posterior axis of the tetrapod; however, a smaller subset of Hox genes is expressed along the anterior-posterior axis of the limbs [89] (or the fins, in the case of gnathostome fish and their pisciform descendants [314]).

Homeobox genes define spatial containers, but not their contents. In hemichordates, the expression of anterior Hox genes leads to downstream development of the kidney and urinary pore in the nose of the organism [230]. In most vertebrates, which share the same Hox genes [330], the kidney is present in the posterior of the organism [132]. Alternatively, the same Hox genes that regulate camera-like eye development in vertebrates also control compound eye development in flies, eyespot development in planarians, and simple-eye development in cnidarians, even though each of the downstream ocular structures has a different form [278].

The Arrangement of Homeobox Gene Expression Defines the Interfaces between Different Spatial Compartments

The arrangement of homeobox genes defines the interfaces between compartments Hox genes control the development of structures at distinct spatial and temporal locations within metazoan embryos. One feature that differs widely between tetrapod embryos is the number of vertebrae. In mice there are 65 segments, in chicks there are 50, in snakes there can be over 300, and in frogs there are ~10 [141]. Humans are born with 33 vertebrae, but adults only have 24, (as 5 are fused to form the sacrum and 4 are fused to create the coccyx) [130]. Both Hox genes and a segmentation clock control vertebrae development (i.e., somatogenesis) [282]. In organisms with more segments, the clock is sped up and somites are produced more rapidly; however, the starting and ending structures in the thoracic region remain a vertebrae regardless of how many segments are present within the region [345]. In pythons the expression of Hox genes also changes the interfaces between regions, eliminating the region where the forelimb would normally exist, and creating a direct connection between the cervical and thoracic vertebrae [90].

The Contents of a Genetically Defined Spatial Compartment Can Be Altered to Provide New Function

In addition to defining interfaces, modularity also simplifies the replacement of parts with those that have a different function in natural biological systems. Limbs in

tetrapods exhibit large differences in morphology, even though the underlying bone structure has a common form. For example the forelimbs of frogs, lizards, birds, humans, cats, whales, and bats are each adapted for different functions, though they share the same bone structure [294]. Attenuation of limbs to rudimentary vestigial structures is another way that body plans can be adapted for new function, by removing structures that are no longer useful. Snakes do not have visible fore or hind limbs, though they are evolved from tetrapods that had both [90]. As described above, the forelimb is eliminated, since the pattern of Hox gene expression no longer allows formation of that region. In the hindlimb region of pythons, however, a vestigial limb is present. Budding of the limb occurs in the embryo, but development does not proceed further. Whales also exhibit a similar vestigial limb structure [51].

Molecular Transport in Multicellular Eukaryotes

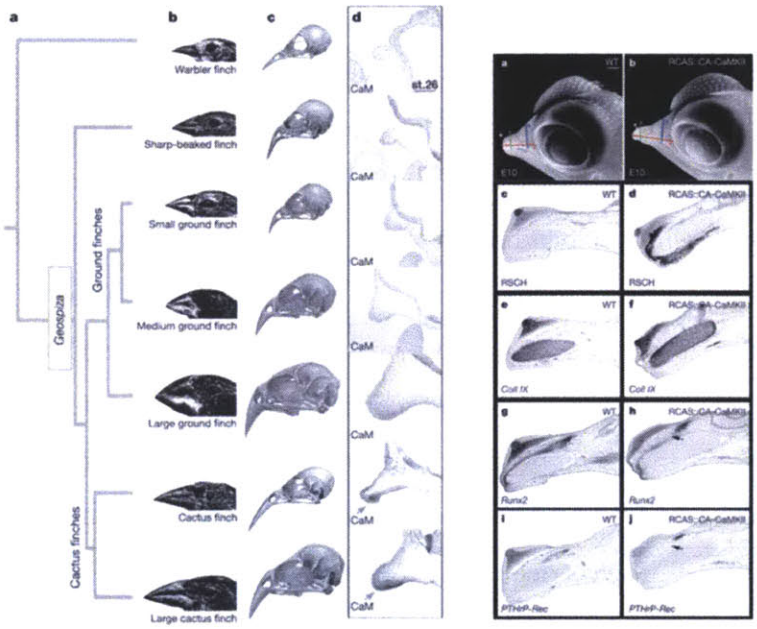
Depending upon their size and complexity, multicellular eukaryotes either actively or passively transport nutrients, wastes and gases between cells. Active transport processes occur in cells, tissues, organs, or organ systems. Cell-cell transport modulators include plasmodesmata in plants and gap junctions in animals or incomplete cell walls in fungi [75]. Simple plants and animals lack vascular systems for nutrient, gas, and waste transfer. Examples of more complex systems for molecular transport include stomata and lenticels for gas transfer in plants, moist skin for gas transfer in earthworms and frogs, spiracles and tracheae in arthropods, gills and circulatory systems in fish, or lung and circulatory systems in tetrapods [75].

1.3.5 Modularity in Engineering and Biological Design

It is remarkable that both nature through evolution and engineers through design create systems using modular components [97]. However, modularity is present in natural systems for the same reasons it is present in engineered systems: modular components are easier to alter and improve while keeping the rest of the system constant. Though engineered systems are typically built for one specific purpose,

biological systems do not have this problem. Natural systems have to be flexible, because natural designs must respond to environmental change. Thus it must be possible to make straightforward changes in biological systems to adapt them to new conditions. In nature these changes occur through evolution across multitudes of organisms where only the winners survive. Architectures amenable to change have emerged because they are more likely to undergo favorable variation and survive [347]. For this reason modularity is present in biological systems at multiple levels [97]. Parts have subparts, which themselves are composed of smaller units. One example of how a natural biological architecture that is amenable to modification is Darwin's finches, whose primary differentiating factor is the size and shape of their beaks. Examination of the factors responsible for the change revealed a single factor (upregulation of calmodulin signaling) resulted in longer and more pointed beaks (see Figure 1-3(a) on the next page) in cactus finches [22]. Alternatively, upregulation of Bmp4 yields broader and deeper beak morphologies in ground finches [21]. Beak size and shape is adapted to match the food supplies present on different islands in the Galapagos. Upregulating calmodulin increased beak length in chicken embryos, whereas increasing Bmp4 increased beak width. That a single genetic switch alters the length along a perpendicular axis and that specific bird types (i.e., ground or cactus finches) tend to vary along that axis means that beak related adaptation to food sources in new environments may require only a simple alteration of one (or maybe two) component(s) . It is intriguing that such natural biological designs facilitate evolution, and can also provide routes to allow engineering when tools are present to modify such parameters.

The similarity between biological and engineered systems leads us to ask the question: can we make biology easier to engineer? Can we implement or use naturally present constraints such as modularity, simplicity, hierarchical design, separation of concerns, abstraction, flexibility, isolation, and standardization to engineer biological systems? The fact that modularity is present throughout biological systems makes it a starting point for engineering. This natural modularity is useful for subsequent engineering because it allows us to think of the different components of the system



(a) Darwin's Finches have Different Levels of Calmodulin Expression and Different Beak Lengths.

(b) Local Upregulation of Calmodulin Increases Beak Length and Pointedness in a Genetically Engineered Chicken.

Figure 1-3: Tuning the Properties of a Modular System. Changes in calmodulin expression changes beak length and shape in both natural (a) and engineered systems (b). The figure is from [22].

independently. In the emerging field of synthetic biology, modularity is a powerful technique because we can either use the existing modularity of systems or engineer them with new modularity. The question here is how can we engineer modularity into biological artifacts or use natural modularity to simplify the implementation of desirable changes.

By engineering modularity into biological systems, we wish to both create new function and construct the systems that we desire more quickly. Rapid construction of systems is an implicit goal of engineering, because it allows one to test and evaluate ideas. The goal is quick and easy design iteration as one progresses towards producing the optimal system. Thus, when engineering systems it is important to be able to do things rapidly. Reflecting on both the Darwin's finch example as well as the broader context of synthetic biology, we wish to both know the changes that are desirable

to make in a biological system and determine an approach to rapidly implement those changes. To do this we need to understand the knobs that we can turn to alter biological designs as well as know how to turn them. Frances Arnold previously stated "There is no such thing as a standard component, because even a standard component works differently depending on the environment. The expectation that you can type in a sequence and can predict what a circuit will do is far from reality and always will be" [279]. However, if we both know which components within a standard system that we should alter and have the technology to adapt standard systems to meet new constraints, then we can begin to engineer biology in a deterministic fashion. History demonstrates that it is possible to create standard parts; engineers previously faced the same challenge in electrical engineering, but were able to create standard components by inventing insulating shielding and other elements that isolated systems from their environment.

Engineering a system requires the composition of small modules into a larger functional system, but it also requires the ability to make sure that a design performs to specification. In engineering biological systems we are still learning how to more quickly and easily build the systems that we desire. In this dissertation we examine three significant examples of modular composition at the protein, DNA, and organismal level.

1.4 Modularity of Natural Biological Systems and Organization of the Dissertation

Modularity is present throughout biological systems. This section describes both natural and engineered modularity in DNA, proteins, and organisms.

1.4.1 DNA

As we describe in Section 1.4.2 on the facing page, DNA contains the information that codes for the expression and sequence of proteins in biological systems. The

information coding for the expression and sequence of proteins is broken up in DNA in a modular fashion. One element encoded by DNA is biological proteins; however, regulatory information is also present that controls under what conditions and how frequently proteins are produced. In prokaryotic systems, the elements leading to expression of a protein include a promoter, a ribosome binding site (RBS), a protein coding region, and potentially a terminator. Sometimes expression cassettes or operons, which contain multiple RBSs with downstream protein coding regions, are present. DNA also is maintained in a modular fashion within cells. In prokaryotic systems, DNA can be present at the plasmid scale (in practice, most commonly 1-15 kb, but can be as large as 150 kb [30]), bacmid scale (typically, 150-350kB [30]), or genome scale (0.6 Mb for *Mycoplasma genitalium* (*M. genitalium*) to 4.6 Mb for *E. coli* [7]). Each of these different types of DNA has biological systems and specific coding regions that ensure its continued maintenance at quantified levels within the cell. In the field of synthetic biology, techniques for synthesizing DNA [198], constructing biological systems from parts [194], modifying DNA [352], and adding new genomic DNA to an organism [137] have been developed.

1.4.2 Proteins

Biological proteins exhibit modularity at multiple levels of structural hierarchy. Protein sequences or primary structure contains information that can be compared with other sequences to predict structures at higher levels of hierarchy [40]. Secondary structures such as alpha helices and beta strands are discrete modular structures within proteins, and the least complex units of structural modularity within proteins [53]. Solving a protein structure reveals the location of these structures relative to the protein sequence, whereas computational analysis predicts their sequence location with high (greater than 80%) accuracy [91]. Tertiary structure is the arrangement of secondary structure into 3D space. Modularity is present at this structural level in within the set of common motifs or folds that emerge through the 3D arrangements of secondary structures in natural systems. The motifs typically found in biological proteins likely are the product of evolution. The 1000 natural protein folds are believed

to have a high "designability" or "number of sequences that possess the structure as their nondegenerate ground state" [219], meaning that secondary structures arranged in these folds can be easily adapted for new purposes through alteration of a few amino acid residues while overall "fold" classification of the structure remains the same. Protein motifs can be partitioned into alpha, beta, alpha plus beta (with distinct alpha helix and beta strand regions), and alpha slash beta structures [162]. In evolutionary history, motifs grew larger and more complex over time, while the alpha slash beta structures appeared most recently [85]. Quaternary structure describes the interaction of proteins with other proteins to form complexes. In this case the proteins themselves are the modular subunits and self assemble through affinity to form larger complexes.

From a functional perspective, protein domains are the least complex modular unit. Protein domains are structural elements that can fold and function independently [200]. Though some proteins contain a single hydrophobic core (i.e., another way to categorize independently folding domains, since the fold occurs around this hydrophobic core [34]), many contain multiple hydrophobic cores in regions that folded individually. As we note below, domains are not necessarily isolated in sequence space (i.e., a domain may contain another internal domain.) Both prokaryotic and eukaryotic proteins contain multiple domains, though multidomain proteins are more common in multicellular organisms. Examples of multidomain proteins found in multicellular eukaryotes include signaling complexes [200] and extracellular mosaic proteins used in cell adhesion [92, 108]. Ponting and Russell, in their review on the "Natural History of Protein Domains" state that there has been "A view of a multidomain protein's function as the sum of its constituent parts [where parts in this context is domains] is obviously simplistic, as it ignores possible interdomain interactions an cooperative effects. Nevertheless this view does provide a first-approximation prediction that is amenable to investigation and subsequent refinement using experimental approaches" [281]. As described above, protein domains can now be identified computationally within proteins using sequence similarity. An interesting feature that displays the intrinsic modularity of protein domains is that the duplication or

insertion of protein domains can often preserve the structural integrity of preexisting protein domains even when it is inserted within an existing domain due to the close proximity of the N- and C-terminus for most domains [281]. Similarities between domains that yield overall function are of interest, as are the elements that yield differences in specificity (such as for substrate choice). Elements that yield differences in specificity provide a focus point for engineering and the development of engineering tools that can modify protein domains in a useful fashion, whereas the elements of commonality between structures may provide a basis for the creation of standard parts at the protein level.

Protein domains as modules and as starting points for standard parts provide a basis for protein engineering using synthetic biology. One possibility is to combine together protein domains to construct new multidomain proteins with novel function. This has been used both for signaling networks [44, 45] and to create composites to optimize synthesis of products [109]. Small target sites within proteins can also be engineered to change the specificity of the protein. Residues of proteins that influence specificity have been previously modified using saturation mutagenesis where one particular region in a protein is targeted [63]. Examination of the 3D structure of a protein often highlights residues that are likely key contributors to function while also identifying elements that contribute to the overall structure but do not directly influence the protein function. For applications such as metabolic engineering, in cases where it is possible to determine protein structures where the substrate is present or absent it may be possible to make simple modifications to increase or reduce the size of or alter a binding pocket optimize an existing reaction or tailor an enzyme for a new reaction [211]. Some synthetic biology tools for constructing multidomain proteins, where flexible linkers connect the protein domains, have been developed. Many of these tools [33, 84, 145, 195, 256, 276] are variants of BioBrick standard biological parts and BioBrick assembly [194] that allow a DNA sequence with the structure of a BioBrick part to be constructed from many smaller BioBrick parts with standard ends in a multistep synthesis reaction. Although the new assembly methods that support protein-fusions are less error prone and faster than standard cloning techniques, they

new assembly techniques are not compatible with the existing collection of over 7,000 BioBrick. In addition, they are not amenable to library creation or the modification of preexisting parts after completion of the assembly process.

1.4.3 Organisms

Group behavior within the organization of cells is controlled by chemical signals between them. The HSLs are used by a large variety of gram-negative bacteria, although a large variety of other signaling molecules exist (e.g., peptides are common in gram-positive bacteria) [309]. Some prokaryotes exhibit additional characteristics of multicellularity, when (often small) aggregates of cells differentiate into different cell types. In these systems, communication between the different cell types is also controlled by chemical signals. As one moves to eukaryotic systems, compartmentalization also occurs internally creating an alternate form of modularity. Early eukaryotes were believed to have absorbed prokaryotes and gained the ability to maintain them internally in simplified forms. Prokaryotes have some analogous systems for compartmentalization either protein (carboxysome and S-layer protein) or lipid based (magnetosome and chlorosome) or lipid, protein, and carbohydrate based (peptidoglycan layer). Both types of organisms have signaling systems that direct resources to these compartments. In addition, complex eukaryotic organisms have complicated differentiation programs, which result in cells that make up different tissues and organs. These functional cell grouping create another layer of modularity within a larger and more complex organism. Now that I have examined natural modularity in three biological systems, I will describe engineered modularity and tools for engineering in the next section (Section 1.6). In synthetic biology, researchers have used chemical signals to spatially control the behavior of organisms [47].

1.5 Abbreviations

The abbreviations used in this thesis are described in the section below. Standard symbols and units appear in an abbreviated for throughout the text.

Chemical Names

3OCHSL		3-oxohexanoyl-homoserine lactone or N-(β-Ketocaproyl)-L-homoserine lactone
DDMA	=	didodecyldimethylammonium bromide
DMPC	=	1,2-Dimyristoyl-*sn*-glycero-3-phosphocholine
DNA	=	deoxyribonucleic acid
EDTA	=	ethylenediaminetetraacetic acid
HSL	=	N-acyl homoserine lactone
IPTG	=	isopropyl β-D-1-thiogalactopyranoside
LB	=	Luria-Bertani, a type media commonly used to culture *E. coli*
M9	=	a minimal media used to grow *E. coli*
Ni-NTA	=	nickel nitrilotriacetic acid
RuBisCO	=	ribulose-1,5-bisphosphate carboxylase/oxygenase

Bacteria and Other Organisms

E. carotovora	=	*Erwinia carotovora*
E. coli	=	*Escherichia coli*
C. crescentus	=	*Caulobacter crescentus*
C. vulgaris	=	*Chlorella vulgaris*
D. radiodurans	=	*Deinococcus radiodurans*
L. sphaericus	=	*Lysinibacillus sphaericus*, formerly known as *Bacillus sphaericus* (*B. sphaericus*)
M. genitalium	=	*Mycoplasma genitalium*
O. vallescia	=	*Ochromonas vallescia*
P. aeruginosa	=	*Pseudomonas aeruginosa*
T. namibiensis	=	*Thiomargarita namibiensis*
S. ureae	=	*Sporosarcina ureae*
S. fragilissima	=	*Syringammina fragilissima*
V. fischeri	=	(*Vibrio fischeri*)

Symbols and Units

2D	=	two-dimensional
3D	=	three-dimensional
Å	=	Ångstroms
°C	=	degrees Celsius
Da	=	Dalton
MW	=	molecular weight

Synthetic Biology Terms

3A	=	three antibiotic assembly
BBa	=	number of a BioBrick assigned by the Registry
BioBrick	=	BioBrick Standard Biological Part
BioBrick assembly	=	BioBrick Assembly Standard for Physical Composition
BioBrick "Prefix"	=	the standard end on the left side of a BioBrick part
BioBrick "Scar"	=	the sequence that forms between two BioBricks during the process of BioBrick assembly
BioBrick "Suffix"	=	the standard end on the right side of a BioBrick part
BioScaffold	=	BioScaffold part
BioScaffold notation	=	written in the form {w,x;y,z}, defines excision site
MAGE	=	multiplex automated genome engineering
"prefix" sequence	=	GAATTCGCGGCCGCTTCTAGAK, where K = G or T or GAATTCGCGGCCGCTTCTAG, when the part starts with ATG and GAATTCGCGGCCGCTTCTAGAG, otherwise
"scar" sequence	=	TACTAGAK, where K=G or T
RBS calculator	=	online system that characterizes expected expression and also can design RBS that wil produce an expected protein level for a specific protein or TACTAG, when the part starts with ATG and TACTAGAG, otherwise
"suffix" sequence	=	always TACTAGTAGCGGCCGCTGCAG

{w,x;y,z}	=	cut sire relative to the BioScaffold, such that w is the position of the left cut on the foward strand, x is the position of the left cut on the reverse strand, y is the position of the righ cut on the forward strand, and z is the position of the right cut on the reverse strand

Biology and Molecular Biology Terms

BSA	=	bovin serum albumin
EPS	=	extracellular polymeric substrate
EYFP	=	enchanced yellow fluorescent protein
GFP	=	green fluorescent protein
GS	=	glycine-serine, a commonly used flexible linker for protein fusions
MBP	=	maltose binding protein
PCR	=	polymerase chain reaction
RFP	=	red fluorescent protein
RBS	=	ribosome binding site
REases	=	restriction endonucleases
RSA	=	a type of S-layer protein created by *Caulobacter crescentus*
S-layer	=	surface-layer, i.e., often refers to the protein that self-assembles into crystals on the outermost layer of Bacteria and Archaea
SDS-PAGE	=	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SbpA	=	a type of S-layer protein created by *L. sphaericus* ATCC 4525 (formerly, *B. sphaericus*)
Type IIB REases	=	restriction enzymes that cleave outside their recognition sequence, Type IIS REases also have this property
USER	=	uracil-specific excision reagent

Electron Microscopy Terms

CCD	=	charge-coupled device
CTF	=	contrast transfer function
EM	=	election microscopy

Containerization Terms

Containerization	=	the process by which shipping containers became standardized
dry-cargo container	=	general purpose container
garmentainers		garment container
GPS	=	global positioning system
RFID	=	radio-frequency identification
tanktainers	=	tank containers

Organizations

ATCC	=	American Type Culture Collection
iGEM	=	International Genetically Engineered Machine Competition
KAUST	=	King Abdullah University of Science and Technology
ISO	=	International Organization for Standardization
MIT	=	Massachusetts Insititute of Technology
NAKFI	=	National Academy of Sciences Keck Futures Intiative
NSF	=	National Science Foundation
Registry	=	Registry of Standard Biological Parts
SynBERC	=	Synthetic Biology Engineering Research Center

Parts and Constructs

BBa_E0040	=	GFP
BBa_G00100	=	verification primer V2 for BioBricks
BBa_G00101)	=	verification primer VF2 for BioBricks
BBa_G1004	=	BioBrick-f, the foward primer for BioBricks
BBa_G1005	=	BioBrick-f, the foward primer for BioBricks
BBa_I7001	=	EYFP
BBa_J70399	=	prototype BioScaffold part, with excision pattern {0,5;15,10}
BBa_J70405	=	composite part formed by the assembly of BBa_R0010

		and BBa_J70399
BBa_J70423	=	composite part formed by assembly of BBa_J70405 and BBa_E0040, test circuit to characterize the prototype BioScaffold
BBa_R0010	=	lacI regulated promoter
pMAL-p5e	=	MBP was amplified from this NEB vector
pSB1AC3	=	high copy BioBrick vector, with ampicillin and chloramphenicol resistance
pSB1AK3	=	high copy BioBrick vector, with ampicillin and kanamycin resistance
pSB1AT3	=	high copy BioBrick vector, with ampicillin and tetracycline resistance
pSB3C5	=	low to medium copy BioBrick vector, with chloramphenicol resistance
pSND-1	=	contains a LuxI gene that encodes for a 3-oxohexanoyl-homoserine (3OC6HSL) synthetase
SbpA	=	a type of S-layer protein created by *L. sphaericus* ATCC 4525 (formerly, *B. sphaericus*)
SbpA-His_7	=	seven-histidine-tagged SbpA containing residues 31-1068
$SbpA_{C-trunc}$-His_7	=	C-terminally truncated SbpA-containing residues 31-918
$SbpA_{N-trunc}$-His_7	=	N-terminally truncated SbpA-containing residues 203-1068
"Sender"	=	contains the plasmid pSND-1 and BBa_I7001 on a pSB3C5 plasmid

1.6 Summary

In this dissertation I focus on three significant examples of building biological systems at the protein, DNA, and organism level. This work, for the most part, derives from

a number of manuscripts where I am the first, or in some cases joint first, author.

Chapter 1 on page 19 examines modularity in biological systems and contains a list of abbreviations for the document.

Chapter 2 on the facing page examines how to engineer several types of biological systems within cells at the DNA level by adding flexibility to modular systems for assembling DNA. This chapter is heavily influenced by the use of evolution in the Belcher laboratory and the invention of BioBrick standard biological parts and BioBrick assembly in the Knight laboratory.

Chapter 3 on page 79 examines the S-layer protein SbpA, a module derived from nature, as a structural unit, both in its monomeric and crystalline state. Upon addition of calcium, SbpA self-assembles into 2D crystals.

Chapter 4 on page 115 examines how to construct systems using clonal cellular populations as modules. In this case we replicate the oxygen profile of a biofilm and examine the responses to of cells to chemical signals.

Chapter 5 on page 137 summarizes the work and explains how modularity can be incorporated into new biological processes.

Chapter 2

BioScaffold Parts for Modular Design of DNA Sequences

In Chapter 1 on page 19, I describe the standard shipping container and the membrane-enclosed vesicle, and how the contents of these modules can be altered without changing the interfaces of the module. I also descibe the structure of bacterial genetic circuits such as operons, which contain a promoter that can change the expression of the proteins in a metabolic pathway or RBSs that can change expression of one protein relative to others. BioScaffold parts provide a way to change the internal contents of a genetic circuit composed of BioBrick parts without changing the external interfaces of the part.

BioBrick assembly is a standard process for composing BioBrick standard biological parts that is analogous to the twistlocks on standard shipping containers, which allow the containers to be joined together using standard connections at any time. The shipping container differs form BioBrick parts in that the twistlocks can be used both to fuse containers together and break them abart, whereas BioBrick parts remain fixed after assembly. In this chapter, I introduce a component called a BioScaffold part that can be assembled together with other BioBricks, but both the sequence of the BioScaffold and (in most cases) some of the surrounding sequence can be altered at a later time point. The flexibility of the BioScaffold endows BioBricks assemblies with a character that is reminiscent of the membrane-enclosed vesicles in eukaryotic

systems. Membrane-enclosed vesicles have attachments and machinery on their outside that controls what things they will attach to; however, the inside can be filled with a desired chemical while the exterior environment remains unchanged. BioBrick parts have standard ends, which allow them to be connected to other BioBrick parts, even if they contain an internal BioScaffold part. Assembly of a BioScaffold part with other BioBrick parts, creates a new BioBrick part but does not prevent the new BioBrick from being assembled together with other BioBrick parts to create larger construsts; however, depending upon the postition of the BioScaffold in a biological circuit, replacement of the BioScaffold part can greatly influence the circuit's functionality.

2.1 Structure and Content

In this chapter, I introduce BioScaffold parts, which are designed to facilitate modular part replacement and functional tuning of genetically engineered systems. BioScaffolds are a family of DNA parts which can be introduced into a DNA sequence, and then can be replaced with alternative parts at a later time. In this chapter, I use BioScaffold parts to engineer protein domains, optimize protein expression levels, and remove BioBrick "scars."

This chapter contains a single manuscript: Julie E. Norville, Ratmir Derda, Saurabh Gupta, Kelly A. Drinkwater, Angela M. Belcher, Andres E Leschziner, and Thomas F. Knight, Jr. "Introduction of customized inserts for streamlined assembly and optimization of BioBrick synthetic genetic circuits" [264]. The original abstract is included in the final dissertation summary in Section 5.1.1 on page 137. Section 1.5 on page 48 contains a list of abbreviations used throughout the dissertation.

2.2 Use of BioScaffolds as a Family of Substitutible Variable Regions That Can Later Be Replaced with Alternative Parts

2.3 Background

In traditional modification of organisms by cloning [237], the emphasis has been on single gene changes that improve the organism or make a single component easier to study. Construction of synthetic genetic circuits brings together many components [118, 137] to accomplish novel tasks, creating functions unobtainable through single gene changes. De novo construction of genetic circuits encompasses the techniques that fall into two categories: techniques for construction and techniques for optimization. Gene synthesis, though decreasing in price [56], still remains prohibitively expensive for de novo synthesis of complete genetic circuits [32]. Instead, either newly synthesized, natural, or existing DNA fragments are pieced together using DNA assembly techniques. A variety of assembly methods now exist including idempotent methods [1, 32, 33, 84, 194, 195, 276, 311], extensions to idempotent methods [52, 83, 145, 256, 262, 338], ligation independent methods [137, 220, 316, 317], Uracil-Specific Excision Reagent (USER) enzyme based methods [72, 114], multipart enzymatic assembly methods [121, 274], and genome-scale assembly methods [106, 138, 168]; however, regardless of how genetic circuits are constructed, novel circuits almost always require modification and optimization. Because understanding of the relevant biological mechanisms remains incomplete [104], one of the main problems in newly assembled circuits is mismatch in the expression levels of the components of the circuit [243]. Optimization, thus, involves modifications of the expression levels [43, 56] to increase desired products, decrease toxic by-products, and increase limiting reagents [161, 251, 365]. A number of existing methods could be used to optimize circuits by rebuilding [104] or reengineering [317, 352] them. It is desired, however, to minimize the number of steps and permit rapid modification

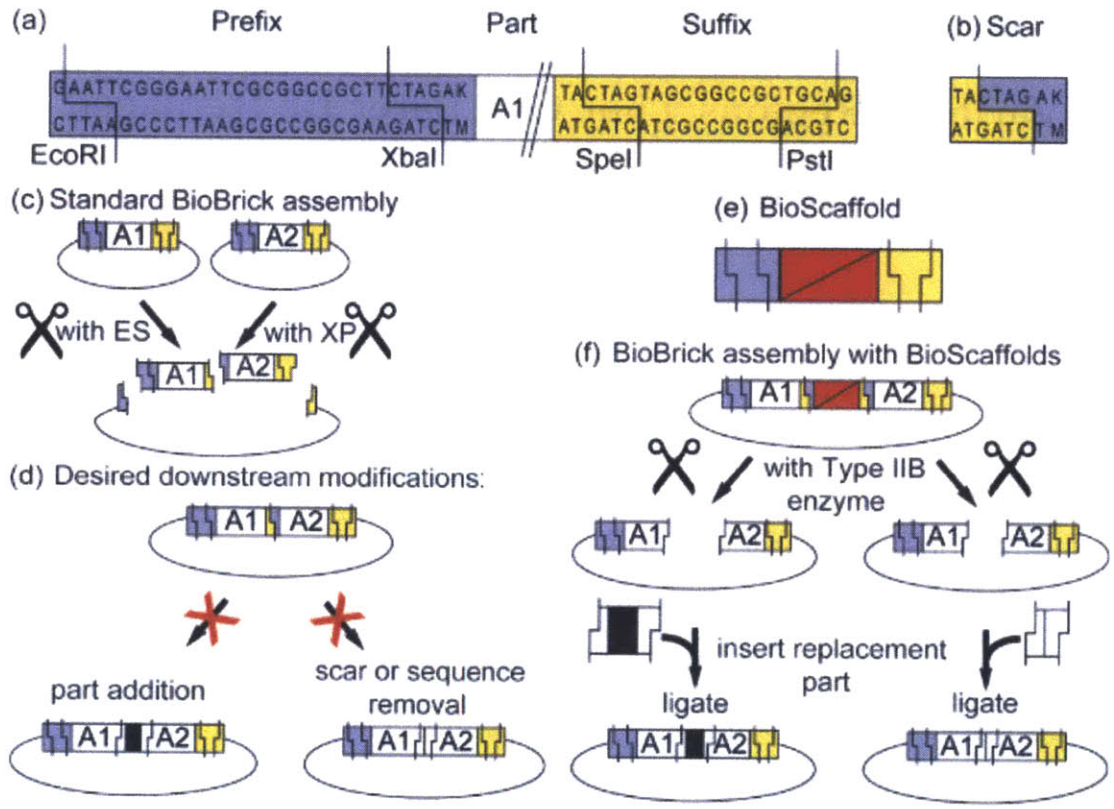


Figure 2-1: Desired BioBrick Circuit Modifications and Approach with BioScaffolds. Every BioBrick standard biological part *(a)* consists of a DNA sequence embedded between a "prefix" sequence (purple box) and a "suffix" sequence (yellow box). Parts may also contain "scars" *(b)*, which form when two parts, such as "A1" and "A2" in *(c)* are fused together using BioBrick assembly [32]. In many cases one would like to convert an undesired scar between two parts in a BioBrick assembly into a different part or completely remove it *(d)*. Our approach is to create a new BioBrick part (the BioScaffold) *(e)*. The BioScaffold can be assembled into a circuit using BioBrick assembly, but unlike normal BioBricks it can be removed and replaced with a new part *(f)*. In this paper we develop a single prototype BioScaffold that illustrates how BioScaffolds can be used to either insert parts or remove scars.

[35, 113, 151, 233, 302, 352, 364]. BioBrick assembly constitutes a widely used strategy for assembly of custom genetic circuits [32]. BioBrick parts are DNA pieces with standard sticky ends (Figure 2-1 c). Using BioBricks, large circuits can be rapidly assembled using a sequence of similar steps. Over 5,000 (now greater than 7,000 in 2012) BioBrick standard biological parts are freely available to researchers through the Registry of Standard Biological Parts [4]. Although BioBricks have been used to construct

a large variety of genetic circuits [3, 25, 26, 49, 55, 70, 76, 152, 163, 190, 215, 239, 311], these circuits often require optimization [28, 38, 99, 190, 333, 337] and currently, there is no standard methodology for optimizing BioBrick circuits.

Modification of a circuit's RBSs is an attractive method for optimization since different strength RBSs create large changes in circuit behavior [352] and a web-based tool is now available to design RBSs of different strengths [302]. Inserting regulatory regions with predetermined sequences changes the levels of protein expression [302, 352] and can have dramatic effects on circuit performance. Optimizing BioBrick circuits using modification of the RBS, however, requires overcoming a fundamental limitation of the BioBrick assembly method: it does not provide a way to insert parts into the circuits once they are assembled (Figure 2-1 on the facing page d). In this paper, we demonstrate a solution to this problem by designing a new BioBrick part, termed BioScaffold (Figures 2-1 on the preceding page e and f), that can be easily excised from intact BioBrick circuits and replaced with other DNA sequences (e.g., RBSs). By virtue of its design, BioScaffolds also bypass two other fundamental limitations of BioBrick assembly: (1) it allows incorporation of parts that contain the sites recognized by the enzymes EcoRI, XbaI, SpeI, and PstI (these sites are incompatible with standard BioBrick assembly) and (2) it allows removal of a scar site (Figures 2-1 on the facing page b and and f) with sequence TACTAGAK (where K = G or T [6]) between BioBrick parts (this is useful because the presence of the stop codon sequence TAG in the scar interferes with production of protein fusions and other modifications [276].)

To outline the design of the BioScaffold part, we first review a hypothetic circuit built from components "A1" and "A2" (designated here as "A1-scar-A2"). Removal of the scar sequence or its replacement with a custom regulatory element (e.g., a RBS) can be performed in two steps. In the first step, excision of the scar and small regions of "A1" and "A2" leaves "sticky ends" inside "A1" and "A2." These ends can then be used for the ligation of the opening with a scar-less DNA sequence or a short sequence that contains a custom RBS. For example, if the excised region is (10 bp of A1)-scar-(12 bp of A2), ligating the gap with 22 base pair annealed repair

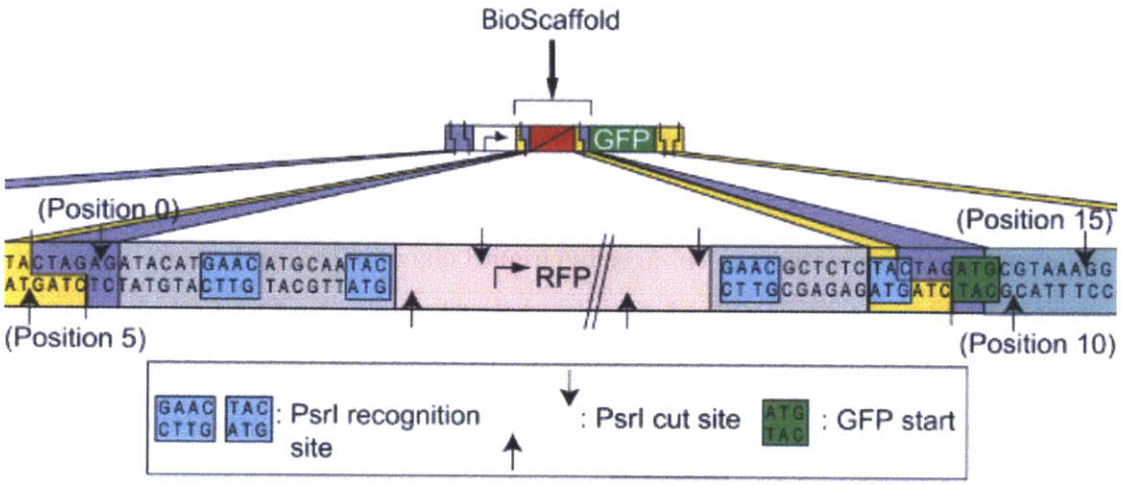


Figure 2-2: The Prototype BioScaffold {0,5;15,10} Embedded in the Test Circuit. Here, the internal structure of the prototype BioScaffold and the locations of the associated PsrI cut sites and recognition sequences are shown. The prototype BioScaffold (part BBa_J70399 in the Registry of Standard Biological Parts) is assembled between a promoter (BBa_R0010, which is not shown here) and GFP (BBa_E0040, which is partially shown here). The BioScaffold contains one PsrI recognition site on either side of the RFP reporter (as well as an internal PsrI site within the RFP reporter that is not shown). PsrI cuts into the scar on the left side of the part and GFP on the right side of the part, allowing the BioScaffold to be replaced with RBS sequences that control the expression of GFP. When the BioScaffold is present, its internal RFP reporter circuit is also present and should produce red fluorescent colonies. The RFP reporter circuit should be removed when the BioScaffold is removed. The cutting profile for the BioScaffold is {0,5;15,10}, using the notation {w,x;y,z} is described in Section 2.4.2 on page 63 of the paper. Different BioScaffolds can be created with different restriction enzyme sites, reporters, or cutting profiles.

oligonucleotides consisting of (10 bp A1)-(12 bp A2) forms a scar-less sequence "A1-A2." While the second step is easily accomplished, the first step requires a specialized enzyme that recognizes scar sequence TACTAGAK and cuts an arbitrary 10 base pairs to the left and 12 base pairs to the right of it. Unfortunately, no enzymes that bind to the sequence TACTAGAK and cut on both sides of the sequence (but outside of it) are available at this time. Restriction endonucleases (REases) that cleave outside their recognition sites are known [149, 238, 296, 297], but none satisfy the specific requirements of this application. The evolution of an enzyme that can excise the scar sequence TACTAGAK is in principle possible [175, 180, 192], but not trivial.

As an alternative, one can use an existing Type IIB or IIS REase that can cleave outside its recognition sequence and modify the "scar region" between parts "A1" and "A2" to introduce the recognition sequence. Table 2.1 on the next page lists an assortment of Type IIB REases that cleave the target on both sides of their recognition sequence [296, 297]. Unfortunately, the cleavage efficiency of most Type IIB REases is low. For example, the efficiencies of ArsI and PsrI are above 56% (e.g., for PsrI more than 70% of DNA fragments can be ligated and 80% of these can be recut), whereas REases commonly used in cloning experiments, such as EcoRI and SpeI, typically have efficiencies above 90% [9]. To facilitate selection of the constructs that will undergo cutting (and subsequent ligation with an arbitrary DNA sequence), we sought to introduce a reporter inside the excised region. Final design of the BioScaffold part, hence, contains two Type IIB recognition sequences placed on either side of RFP reporter (as well as an additional site within the RFP reporter) that serves as a selection marker (Figure 2-2 on the facing page). This configuration makes it simple to select colonies that circuits in which the BioScaffold has been placed or excised and replaced with repair oligonucleotides (Figure 2-3 on page 64).

In this paper we demonstrate the utility of a BioScaffold to optimize BioBrick circuits (by inserting a series of RBS regulatory sequences) as well as for production of protein fusions. Because none of these processes can be easily attained through use of standard BioBrick assembly, these results demonstrate that the use of BioScaffolds can aid in overcoming several limitations of BioBrick assembly.

2.4 Results

2.4.1 Maximum Excision Capacity of Commercially Available Type IIB REases

As described in the preceding section (Section 2.3 on page 57), the prototype BioScaffold is primarily useful for introducing RBSs (Circuit Tuning BioScaffold) or N-terminal protein tags (Protein Engineering BioScaffold) into BioBrick circuits that

Table 2.1: BioScaffold Designs for Maximal Excision. Several Type IIB enzyme recognition sites are aligned to the scar sequence TACTAGAK to determine maximum excision to the left and to the right of the BioScaffold. The alignment of the recognition sites to the scar fixes the sequence at the start and end of the BioScafffold. We include several notes to clarify the table. First, the enzymes shown cut on both sides of their recognition site, not just one. For example, the cut sites and recognition sequence for PsrI is (7/12)GAACNNNNNNTAC(12/7) [297]. Second, the K (in the scar sequence TACTAGAK) is T for any protein coding region or other sequence that starts with ATG (i.e., TACTAGATG) and G for any other sequence. M represents A or C, R represents G or A, and Y represents C or T [6]. Third, recognition sequences for the enzyme are highlighted in *bold font.* Fourth, the internal cuts sites within the BioScaffold are not shown and the selection marker between the two recognition sites is represented as ... In the prototype BioScaffold, the selection marker is a RFP reporter circuit. Fifth, the notation represents the location of the cut sites in condensed form. BioScaffold {w,x;y,z} notation is described in Section 2.4.2 on the next page of the paper.

Enzyme	Cut site and upstream scar	BioScaffold start	Selection Marker	BioScaffold end	Downstream scar and cut site	Notation
AlfI	↓NNNTACTAGAG ↑NNNNNATGATCTC	**CANNNNNNTGC** **GTNNNNNNACG**	—....—	**GCANNNNNNTGC** **CGTNNNNNNACG**	TACTAGAKNNNN↓ ATGATCTMNN↑	{10,12;12,10}
ArsI	↓NNTACTAGAG ↑NNNNNNATGATCTC	**ACNNNNNNTTYG** **TGNNNNNNAARC**	—....—	**CRAANNNNNNGTC** **GYTTNNNNNNCAG**	TACTAGAKNNNNN↓ ATGATCTM↑	{8,13;13,8}
BaeI	↓NNNNNNNNNTACTAGAG ↑NNNNNNNNNNNNNNATGATCTC	**TAYC** **ATRG**	—....—	**GRTACNNNNG** **CYATGNNNNC**	TACTAGAKNNNNNNNN↓ ATGATCTMNNN↑	{16,21;16,11}
BarI	↓TACTAGAG ↑NNNNNATGATCTC	**AAGNNNNNNTAC** **TTCNNNNNNATG**	—....—	**GAAGNNNNNN** **CTTCNNNNNN**	**TAC**TAGAKNNNNNNN↓ **ATG**ATCTMNN↑	{7,12;15,10}
BcgI	↓NNNTACTAGAG ↑NNNNNATGATCTC	**CANNNNNNTCG** **GTNNNNNNAGC**	—....—	**GCANNNNNNTCG** **CGTNNNNNNAGC**	TACTAGAKNNNN↓ ATGATCTMNN↑	{10,12;12,10}
BdaI	↓NNTACTAGAG ↑NNNNATGATCTC	**TGANNNNNNTCA** **ACTNNNNNNAGT**	—....—	**TGANNNNNNTCA** **ACTNNNNNNAGT**	TACTAGAKNNNN↓ ATGATCTMNN↑	{9,11;12,10}
BplI	↓NNNTACTAGAG ↑NNNNNNNNATGATCTC	**NNNNNCTC** **NNNNNGAG**	—....—	**GAGNNNNNCTC** **CTCNNNNNGAG**	TACTAGAKNNNNN↓ ATGATCTM↑	{10,15;13,8}
BsaXI	↓NNNNNNNNTACTAGAG ↑NNNNNNNNNNNNATGATCTC	**CTCC** **GAGG**	—....—	**GGAGNNNNNG** **CCTCNNNNNC**	**TAC**TAGAKNNNNN↓ **ATG**ATCTMNN↑	{15,18;13,10}
CspCI	↓NNNTACTAGAG ↑NNNNNATGATCTC	**CAANNNNNGTGG** **GTTNNNNNCACC**	—....—	**CCACNNNNNTTG** **GGTGNNNNNAAC**	TACTAGAKNNNNN↓ ATGATCTMNNN↑	{10,12;13,11}
FalI	↓TACTAGAG ↑NNNNNATGATCTC	**AAGNNNNNCTT** **TTCNNNNNGAA**	—....—	**AAGNNNNNCT** **TTCNNNNNGA**	**TAC**TAGAKNNNNNN↓ **ATG**ATCTMN↑	{7,12;14,9}
PpiI	↓NNNTACTAGAG ↑NNNNNNNNATGATCTC	**NNNNNCTTC** **NNNNNGAAG**	—....—	**GAACNNNNNCTC** **CTTGNNNNNGAG**	TACTAGAKNNNNN↓ ATGATCTM↑	{10,15;13,8}
PsrI	↓TACTAGAG ↑NNNNNATGATCTC	**AACNNNNNNNTAC** **TTGNNNNNNNATG**	—....—	**GAACNNNNNN** **CTTGNNNNNN**	**TAC**TAGAKNNNNNNN↓ **ATG**ATCTMNN↑	{7,12;15,10}
TstI	↓TACTAGAG ↑NNNNNATGATCTC	**GANNNNNNGTG** **CTNNNNNNCAC**	—....—	**GGANNNNNNGTG** **CCTNNNNNNCAC**	TACTAGAKNNNNN↓ ATGATCTM↑	{7,12;13,8}

do not contain internal PsrI restriction enzyme recognition sites or RFP reporters. Thus, a variety of BioScaffolds might be desirable for other applications or for use with BioBrick circuits that contain PsrI sites or RFP reporters. To provide a sense of the extendibility of BioScaffolds, we examined a variety of restriction enzymes that can cut outside their restriction enzyme recognition sites and can thus be used to create the cloning site within a BioScaffold. For several commercially available Type IIB enzymes we have determined the maximum number of nucleotides from the surrounding parts "A1" and "A2" that can be excised using an intermediate BioScaffold inserted between them. To find the maximum number of excised nucleotides for each enzyme, we align the enzyme recognition sites to the outermost position where they can bind to the left and right scar sequences. The resulting alignments, shown in Table 2.1 on the preceding page, demonstrate that the furthest distances the enzymes can cut into left part "A1" and the right part "A2" are 9 nucleotides into "A1" for enzyme BaeI and 3 nucleotides (plus 2 additional nucleotides for downstream protein parts) into "A2" for enzymes BaeI and CspCI. When part "A2" is a protein coding region then the last 2 nucleotides from the scar region are included in part "A2," since the last two nucleotides of the scar TACTAGAK become the first 2 nucleotides of the start codon ATG.

2.4.2 BioScaffold Notation, {w,x;y,z}

Since other BioScaffolds beyond the circuit tuning and N-terminal protein engineering application (or for the same function (tuning circuits) where PsrI restriction enzyme sites or RFP reporters are present) might be desired, we created a notation to make it simpler to compare two BioScaffolds. We define the following notation {w,x;y,z} to describe in condensed form the positions of the cut sites relative to the BioScaffold: w is the position of the left cut on the forward strand, x is the position of the left cut on the reverse stand, y is the position of the right cut on the forward strand, and z is the position of right cut on the reverse strand. The position numbering begins on the right side of the BioScaffold part with 0 as the first position before TACTAGAK, where numbers increase to the right, and on the left side with -1 as

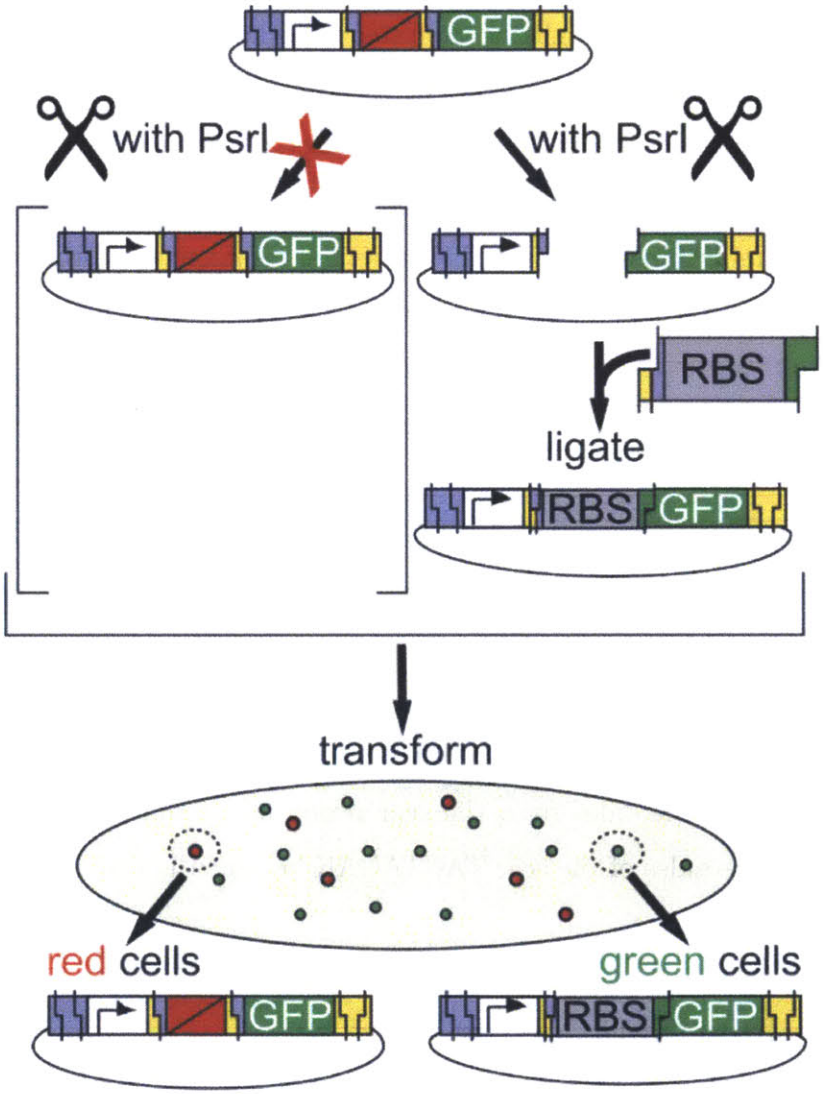


Figure 2-3: Excision and Selection for BioScaffold Removal. Here, the prototype BioScaffold is present in a test circuit, with sequence "promoter-scar-BioScaffold part-scar-GFP." It is cut with the enzyme PsrI and then replaced with a RBS sequence. The BioScaffold test part contains a RFP expression circuit surrounded by two PsrI recognition sites; thus, cells that contain the BioScaffold part, such as the cells that contain the test circuit, exhibit red fluorescence. If PsrI excises the BioScaffold part from the test circuit and an RBS is ligated across the open gap, then the sequence "promoter-scar-RBS-GFP" is obtained. Cells that contain this sequence exhibit green flourescence, where the strength of the fluorescence depends on the RBS used. To show the flexibility of the BioScaffold, it was also replaced with a "RBS-MBP-GS" sequence to create the sequence "promoter-scar-RBS-MBP-GS."

the first position after the K of TACTAGAK, where numbers increase to the left– see Table 2.1 on page 62 for examples. The cut sites of the prototype BioScaffold are given by {0,5;15,10} (see Figure 2-2 on page 60 where the excision positions are marked above the cut sites). When referring to a BioScaffold in BioBrick format, the notation {w,x;y,z} assumes that the BioScaffold is in its most common form when used with BioBricks (i.e., surrounded by a scar sequence on either side).

2.4.3 Testing the BioScaffold

As a demonstration, we positioned the prototype BioScaffold in a synthetic circuit between a promoter and GFP (Figure 2-3 on the facing page). Replacement of the BioScaffold with an RBS caused this test circuit to become a GFP reporter (containing a promoter, RBS, and GFP) that expresses GFP within the cell (Figure 2-4 on the next page). Alternatively, replacement of the BioScaffold with a RBS-(maltose-binding protein)-(glycine-serine) (RBS-MBP-GS) sequence created a circuit that produced a MBP-GS-GFP fusion protein (Figure 2-5 on page 68) that is fluorescent green and binds to amylose resin. The prototype BioScaffold has been designed to contain specific cut locations on either side of its sequence (Figure 2-2 on page 60). After excision (Figure 2-3 on the facing page), the BioScaffold will be replaced with one of several RBSs that are designed to drive well-defined levels of expression in the downstream gene, demonstrating how BioScaffolds can be used to facilitate the optimization of circuits. Alternatively, replacement of the BioScaffold with the MBP-GS fusion protein part will demonstrate how BioScaffolds can be used to create protein fusions. The visual markers used in the prototype system help track the presence of the BioScaffold (which contains an RFP reporter) and the performance of the optimization process (effect of different strength RBSs or a MBP-GS fusion on the GFP reporter). Figure 2-3 on the preceding page shows graphically the proposed replacement of the BioScaffold with repair oligonucleotides that contain an RBS sequence, which will affect the performance of the final GFP reporter circuit (Figure 2-4 on the following page). Figure 2-5 on page 68 shows that the BioScaffold can successfully be replaced with an expressed MBP-GS fusion, which causes the

final protein fusion to gain the property of affinity to an amylose column as well as maintaining the fluorescence of GFP.

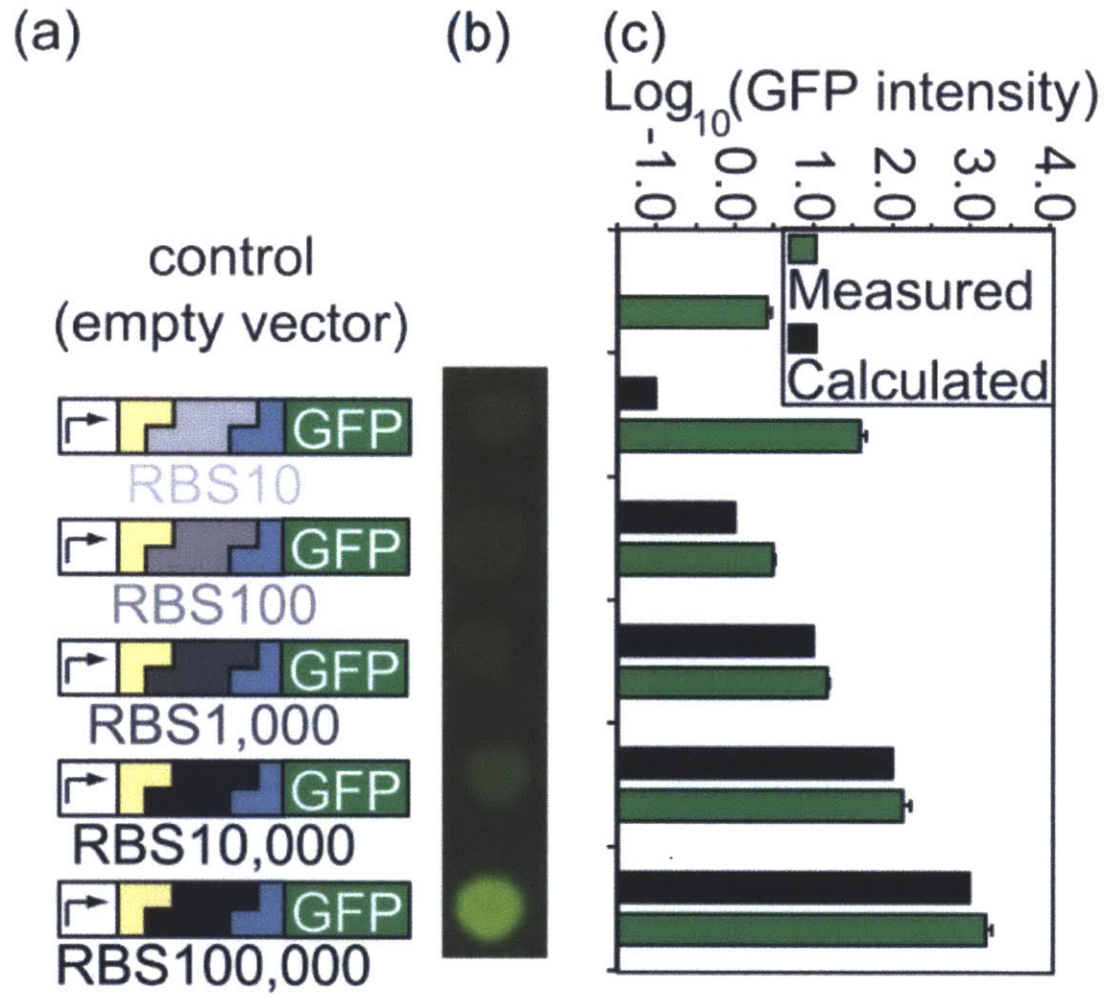


Figure 2-4: GFP Expression Levels Created by the Different RBSs. Sequence verified clones of the form "promoter-scar-RBS-GFP" *(a)* were analyzed visually *(b)* and using flow cytometry *(c)*. The inserted RBSs were expected to drive relative translation initiation rates of 10, 100, 1,000, 10,000, and 100,000 for the downstream sequence GFP with the cell fluorescence proportional to the translation initiation rate *(a)*. Qualitative visual assessment revealed green fluorescent intensity commensurate with the expected values *(b)*, although colony thickness can influence perceived intensity. Simultaneously transformed colonies of RBS10, RBS100, and RBS1,000 appeared light green, while RBS10,000 appeared green and RBS100,000 appeared bright green *(b)*. Quantitative assessment using flow cytometry data revealed GFP intensity levels commensurate with expected values, except for the higher than expected translation initiation of GFP driven by RBS10 *(c)*.

2.4.4 Excision of the BioScaffold and Replacement with the RBS and the RBS-MBP-GS Sequences

Excision of the BioScaffold from the "promoter-scar-BioScaffold-scar-GFP" composite part and ligation of the RBS insert was expected to create a "promoter-scar-RBS-GFP" composite (where one RBS was inserted in each of five parallel reactions). The plated transformations of each reaction contained a mixture of red colonies and green colonies. It was assumed that the red colonies resulted from clones where the BioScaffold was not excised. At the optimal concentration of vector in the cutting reaction, less than 3% per 1,000 colonies were red (i.e., still contained the BioScaffold.) We picked 25 non-red clones for sequencing (5 for each RBS insert). For 80% of the non-red clones, sequencing showed that the RBS was inserted in place of the BioScaffold part at the correct location with the correct sequence. For 12% of the clones, sequencing showed that the entire test circuit was completely mutated out. For 8% of the clones, the sequencing result was noisy, preventing interpretation. We speculate that none of the errors are due to the cutting properties of the enzyme, but rather to selection against the circuit or inadvertent picking of multiple colonies. Using the same starting circuit, selection, and circuit verification protocol, ligation of the RBS-MBP-GS insert was expected to create a "promoter-scar-MBP-GS-GFP" circuit where the ATG start codon of the GFP was removed by the BioScaffold.

2.4.5 Expression Levels of the RBSs

Colony and sequence verified clones that contain RBSs in place of the BioScaffold part were assessed qualitatively and quantitatively (Figure 2-4 on the facing page). Visual analysis demonstrated that the clones exhibited a range of green fluorescent intensities: the lowest predicted expression levels appeared light green and the highest expression levels appeared bright green (Figure 2-4 on the preceding page). Visual analysis is only a rough guide to intensity since the number of cells present in a colony could influence its appearance. Overall, the flow cytometry data confirmed the results from [302] using our method to introduce the RBS. Flow cytometry data

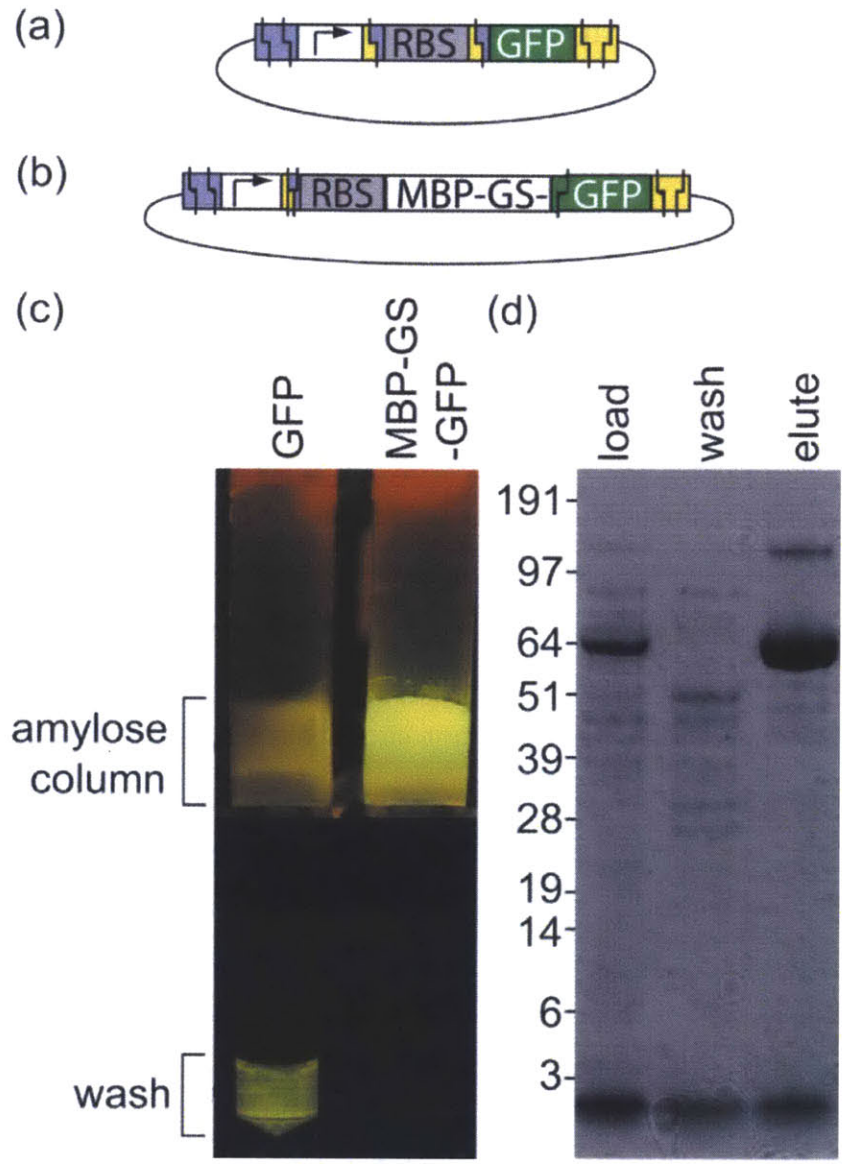


Figure 2-5: Properties of GFP with an N-Terminal MBP-GS Fusion. Lysate supernatants from cultures expressing GFP alone *(a)* or the MBP-GS-GFP fusion *(b)* were applied to amylose columns. The green fluorescent product was not retained on the "GFP" column when column buffer was applied, but rather appeared in the first wash eluate *(c)*. In contrast the green fluorescent product was retained on the "MBP-GS-GFP" column when column buffer was applied *(c)*. The protein purification process for MBP-GS-GFP was monitored using SDS-PAGE: Lane 1, cell lysate ("load"); Lane 2, first wash eluate ("wash"); Lane 3, maltose eluate ("elute") *(d)*. Mass spectroscopy confirmed that the band at ~67 kDa was MBP-GS-GFP and identified the minor contaminant at the top of Lane 3 ("elute") as a dimer form. The dimer is known to be a normal occurrence with this construct [12, 82].

indicated that the predicted expression levels were commensurate with the actual intensity values except for sequence RBS10, which was expected to produce the lowest intensity of protein fluorescence (Figure 2-4 on page 66). The actual expression level for RBS10 was higher than expected; however, similar deviations are observed by Salis, et al. [302] for low RBS translation initiation strengths, and therefore appear to be unrelated to the use of the BioScaffolds. If it was important to create a circuit with low expression of GFP, we could have performed another optimization round using RBSs with similar but slightly different relative translation rates, such as 8, 9, 11, and 12.

2.4.6 Affinity of the MBP-GS-GFP Protein to an Amylose Column

The following observation confirmed the presence of desired genetic product (MBP-GS-GFP) in the cells that contained the "promoter-scar-RBS-MBP-GS-GFP" circuit (Figure 2-5 on the preceding page a): (1) cells which contained MBP-GS-GFP exhibited strong fluorescence similar to those of cells that contained the classic BBa_J04430 GFP reporter circuit [2] (Figure 2-5 on the facing pageb), (2) running a lysate from a culture of MBP-GS-GFP cells through an amylose column led to retention of the fluorescent product, which could only be eluted with a buffer containing maltose, whereas for GFP cells the fluorescent product was eluted in the wash (Figure 2-5 on the preceding page c) and (3) both crude MBP-GS-GFP lysate and eluate from the amylose column contained a strong band at ~67 kDa and a weak band at ~134 kDa in a SDS page gel (Figure 2-5 on the facing page d, "load" and "elute" lanes), which are the expected sizes of a MBP-GS-GFP fusion and a MBP-GS-GFP dimer [12, 82], rather than the 40 kDa length of MBP-GS or the 27 kDa length of GFP. The identity of the bands was confirmed by mass spectrometry. (The concentric patterns seen on the gel are artifacts that come from the glass of the scanner.)

2.5 Discussion

2.5.1 Use of the BioScaffold Test Circuit to Overcome Limitations in BioBrick Assembly and Comparison with Other Methods

The original protocol for BioBrick assembly does not allow insertion of parts into premade assemblies and optimization of the performance of premade circuits. We demonstrate that BioScaffolds enable optimization of premade circuits through insertion of RBSs of predetermined strength and protein fusions. Although a few other approaches have been recently developed for the same purpose (see below), the use of BioBricks provides advantages and flexibility not present in these methods. For example, Church and co-workers demonstrated rapid evolution of circuits with RBSs using Multiplex Automated Genome Engineering (MAGE) to create and select for circuits with improved performance [352]. Sauro and co-workers used the Clontech In-Fusion PCR Cloning kit to re-engineer BioBrick circuits [317] after assembly. Substitution of the BioScaffold with RBSs provides an approach different from MAGE and In-Fusion because the expression of proteins was manipulated rationally using the RBS calculator [302]. In addition, BioScaffolds can be used to insert both long and short sequences.

BioBrick parts cannot contain restriction sites for EcoRI, SpeI, XbaI, and PstI; thus, conversion of arbitrary DNA sequences into BioBrick parts requires the time-consuming and expensive process of removing these sites from the sequence of interest. We demonstrate, however, that this BioScaffold can be used to insert any parts into BioBrick circuits except for those that contain PsrI recognition sites. Importantly, PsrI recognizes GAACNNNNNNTAC (or GTANNNNNNGTTC) and the probability of encountering this sequence in a random DNA sequence (once every 8,192 base pairs for the forward or reverse recognition site) is 8 times lower than that of encountering a site for EcoRI, SpeI, XbaI, or PstI. In the event that the site for PsrI is present in the insert part, it is possible to use another restriction enzyme (see Table 2.1 on

page 62) to create appropriate sticky ends for the insert.

2.6 Conclusions

We demonstrated that BioScaffolds could be implemented as a BioBrick part, integrated into BioBrick circuits, and used to remove a scar sequence. Additionally, BioScaffolds allowed the introduction of parts directly into preassembled circuits. This work demonstrated the introduction of RBSs for circuit optimization, but natural sequences, BioBrick parts, or assemblies of parts can alternately be incorporated. Specifically, appropriately designed BioScaffolds allow the creation of protein fusions or the addition of N- or C-terminal tags. For instance, the prototype BioScaffold shown here can also be used to add N-terminal tags, as was demonstrated here through the introduction of a MBP fusion upstream of GFP. Another advantage of BioScaffolds is that a wide variety of parts can be introduced at a specific position in a single reaction, providing an approach to perform directed evolution and selection of circuits using BioBrick standard biological parts. Thus, BioScaffolds provide a compelling tool to extend idempotent assembly techniques, such as BioBrick assembly, and can even be utilized in combination with PCR based circuit assembly techniques.

2.7 Methods

2.7.1 Construction of Prototype BioScaffold: Part BBa_J70399

The prototype BioScaffold (BBa_J70399, the part number assigned by the Registry of Standard Biological Parts at [4]) was created using the RFP production circuit (BBa_J04450) as a template and the primers J70399-f (5'-gtttcttcgaattcgcggccgcttctag agatacatgaacatgcaatacgcaaacc-3') and J70399-r (5'-gtttcttcctgcagcggccgctactagtagag agcgttcaccgacaaacaacag-3'). Each primer contains a recognition site for the Type IIB REase PsrI as well as the standard BioBrick ends. The reactions contained 45 µl PCR SuperMix High Fidelity (Invitrogen), 12.5 pmoles of forward and reverse primer, and 1 ng template DNA in a 50 µl total volume. The PCR steps included a

denaturation step of 96°C for 4 minutes followed by 36 cycles of a 94°C denaturation step for 30 seconds, a 52.3°C anneal step for 30 seconds, and a 68°C extend step for 2.5 minutes. Finally, the reactions were incubated at 68°C for 10 minutes before being cooled to 4°C until the reactions were halted. The samples and 1 μg of 2-log DNA ladder (New England Biolabs, Inc.) were electrophoresed in separate lanes on a 1% agarose gel. Sample bands of length 1000 base pairs were excised and purified with a QIAEX II Gel Extraction Kit (QIAGEN). The amplified linear DNA fragment was cloned into the BioBrick vector pSB1AT3, by digesting both the fragment and the vector with XbaI and PstI and performing the ligation using protocols adapted from BioBricks assembly kit (New England Biolabs, Inc.) to place the fragment into the vector. The ligation mixture was transformed into chemically competent *E. coli* strain TOP10 (Invitrogen) [311] and plated on LB agar plates supplemented with 15 μg/ml tetracycline and 100 μg/ml ampicillin.

2.7.2 Assembly of the BioScaffold Test Circuit: Part BBa_J70423

The test circuit consists of a promoter, the BioScaffold, and GFP assembled into the circuit "promoter-scar-BioScaffold-scar-GFP." The assembly was performed in two rounds using the BioBrick Assembly kit (New England Biolabs, Inc.) for three antibiotic (3A) assembly. First, the promoter (BBa_R0010 in BioBrick vector pSB1A2), the BioScaffold part (BBa_J70399 in BioBrick vector pSB1AT3), and the destination vector pSB1AK3 were digested. The upstream part, the promoter (BBa_R0010), was digested with EcoRI and SpeI. The downstream part, the BioScaffold part (BBa_J70399), was digested with XbaI and PstI. The destination vector pSB1AK3 was disgested with EcoRI and PstI. The digests were mixed together and ligated with T4 ligase, to form a composite part (BBa_J70405 in the vector pSB1AK3). The ligation mixture was transformed into chemically competent TOP10 cells as above and plated on LB agar plates containing 100 μg/ml ampicillin and 50 μg/ml kanamycin. 5 colonies that appeared red were verified by colony PCR and sequencing as described below.

Next, the composite part (BBa_J70405 in BioBrick vector pSB1AK3), the GFP (BBa_E0040 in BioBrick vector pSB1A2), and the destination plasmid pSB1AC3

were digested and ligated in the same manner as above, where BBa_J70405 was the upstream part and BBa_E0040 was the downstream part, to form the test circuit composite part (BBa_J70423). The ligation mixture was transformed into chemically competent TOP10 cells as above and plated on LB agar plates containing 100 μg/ml ampicillin and 35 μg/ml chloramphenicol. 5 colonies that appeared red were verified by colony PCR and sequencing as described below.

Table 2.2: Sequences Generated by the RBS Calculator. 10, 100, 1,000, 10,000, and 100,000 were entered as target initiation rates in the RBS calculator. The RBS calculator uses a relative scale to relate relative translation initiation rate with five terms that quantify the strengths of molecular interactions involved in this process, with the result that any RBS sequence designed by the RBS Calculator can be related on this scale (further descriptions of this unitless parameter are available in the document [302] and FAQs [16, 17] describing this software. Qualitatively from visual analysis of GFP expression in our samples and quantitatively based on our FACS results, we consider 10-100 to be low rates of translation initiation, 1,000 to be a medium rate, and 10,000 and above to be high rates of translation initiation.

Target	Rate	Sequence
10	10.5	CTAAATAGGAGGCTGGGAGTTCAACGAAACCCCT
100	95.6	CCCCCGTTCACTATACCGCAGGCCTTCTTTACAAA
1,000	950.1	ACATTAACCTACAAAGAACGTCGCAGAGGGA
10,000	9,953.7	TGTCGCGGATACTGATCCATAAAGGCCGGGGTT
100,000	94,459.0	TAGAGCCGTTAAAGAAGCTAGGAGGCCGAA

2.7.3 Design of RBSs

The RBS calculator [15] was used to design the RBS sequences. The downstream GFP (BBa_E0040) was entered as the protein coding sequence. TACTAGAG was used as the presequence. Five different target translation initiation rates (10, 100, 1,000, 10,000, and 100,000) were entered into the calculator to generate the five different RBS sequences. The sequences generated by the calculator (see Table 2.2) were then utilized to create forward and reverse oligonucleotides, taking into account the cutting profile of the BioScaffold. The forward oligonucleotide has the sequence "G-RBS sequence-ATGCGTAAA" and the reverse oligonucleotide has the sequence reverse complement of "CTAGAG-RBS sequence-ATGC." After the replacement of

the BioScaffold with the annealed oligonucleotides, the downstream scar disappears, while the upstream scar remains.

2.7.4 Design of the RBS-MBP-GS Insert

The RBS-MBP-GS insert was created using the plasmid pMAL-p5e (New England Biolabs) as a template and the primers RBS-MBP-GS-GFPstart-f (5'-gaattcgcggccgcttctag agacgaacgctctctactagatgcctagagtcgccccctaagggcggaggtaggagaaactcaaatgaaaatcgaagaag gtaaactggtaatctg -3') and RBS-MBP-GS-GFPstart-r (5'-ctgcagcggccgctactagtagtaatata tgttcgatagattttacgagaaccagtctgcgcgtctttcagg -3') using the same PCR and gel extraction protocol as for BBa_J70399.

2.7.5 Excision of BioScaffold Part from the Test Circuit and Replacement with RBSs

The oligonucleotides containing the RBSs were prepared by combining 8 μl of 100 μM forward oligonucleotide, 8 μl of 100 μM reverse oligonucleotide, 10 μl annealing buffer (100 mM Tris HCl pH 7.5, 1 M NaCl, 10 mM EDTA), and 74 μl milliQ water. They were then annealed by heating to 95°C for 2 minutes, ramping from 95°C to 25°C over 45 minutes, and then cooling to 4°C. We diluted 10 μl of oligonucleotides into a final volume of 1000 μl.

We digested the test circuit by combining 100 ng DNA, 1x Buffer Y (SibEnzyme, Inc.), 100 μg/ml Bovine Serum Albumin (BSA) (SibEnzyme, Inc.), and 0.5 μL PsrI (SibEnzyme, Inc.) into a final volume of 50 μl. The restriction digest reaction was incubated for 1 h at 30°C followed by 20 minutes at 65°C. In our experience digestion of more DNA dramatically increases the number of undigested products yielding red colonies upon transformation.

We performed the ligation by combining 5 μL of the PsrI digestion reaction (10 ng DNA), 0.2 μL of the diluted annealed oligonucleotides, 1x T4 DNA ligase reaction buffer, 200 units of T4 DNA ligase into a 20 μL total volume and cooling to 18°C for 30 minutes. The ligation mixture was transformed into chemically competent TOP10

cells as above and plated on LB agar plates containing100 μg/ml ampicillin and 35 μg/ml chloramphenicol. 5 non-red colonies per reaction were verified by colony PCR and sequencing as described below.

2.7.6 Replacement of the BioScaffold with the RBS-MBP-GS Insert to Create BBa_J70631

Both the test circuit and the gel purified RBS-MBP-GS insert were digested with PsrI as described above in separate reactions. We performed the ligation by combining 5 μl of the test circuit digestion reaction, 5 μl of the RBS-MBP-GS insert digestion reaction, 1x T4 DNA ligase reaction buffer, 200 units of T4 DNA ligase into a 20 μl total volume and cooling to 18°C for 30 minutes. The transformation and sequencing of BBa_J70631 occurred as described previously for a test circuit containing an insert.

2.7.7 Verification with Colony PCR and Sequencing

Colony PCR and subsequent gel electrophoresis were performed according to [311] except that the BioBrick primers BioBrick-f (BBa_G1004) and BioBrick-r (BBa_G1005) were used. The PCR protocol was the same as described for the BioScaffold part above, except that the extension step at 68°C occurred for 3.5 minutes instead of 2.5 minutes. The Massachusetts Institute of Technology Biopolymers Laboratory performed DNA sequencing, using the verification primers VF2 (BBa_G00100) and VR (BBa_G00101).

2.7.8 Flow Cytometry Measurement of Final Circuit Fluorescent Intensity

Expression data were collected using a Becton-Dickinson FACSCAN flow cytometer with a 488-nm argon excitation laser and a 515-545 excitation filter [131]. Cells were grown in M9 media [65] (1x M9 salts, 2 mM MgSO4, 0.5% glycerol, 0.2% casamino acids, 2 mM thiamine) with ampicillin and chloramphenicol. The cultures were grown

according to [29], except that the samples were measured in M9 medium [65].

2.7.9 Purification of the MBP-GS-GFP Protein Using an Amylose Column

The BBa_J70631 construct, which contains a MBP-GS-GFP expression circuit, was expressed in *E. coli* Top10. The culture was grown overnight at 37°C in rich media (10 g typtone, 5 g yeast extract, 5 g NaCl, 2 g glucose) containing 100 μg/ml ampicillin and 35 μg/ml chloramphenicol, expanded 1:100 the next day and grown for 3 h after the OD600 reached 0.6. The BBa_J04430 construct, which contains a GFP expression circuit, was grown in the same manner except the media did not contain chloramphenicol. Purification of a 40 ml culture of each construct was performed according to [295] using amylose resin (New England Biolabs), expect that the GFP and MBP-GS-GFP lysates were diluted until they exhibited equivalent fluorescent intensity and an equivalent volume of each cleared lysate (containing less MBP than the binding capacity of the amylose resin) was applied to each column.

2.7.10 SDS-PAGE of MBP-GS-GFP

Samples were run on a 4-12% Bis-Tris gel (Invitrogen) and stained with Coomassie blue or SimplyBlue (Invitrogen) stain [189]. The contrast of the image was adjusted uniformly in Adobe Photoshop to simplify the visualization of the bands.

2.7.11 Mass Spectrometry of MBP-GS-GFP

The lanes from a SimplyBlue stained SDS PAGE gel were excised and submitted for analysis to the Proteomics Core Facility of the Koch Institute for Integrative Cancer Research at MIT. The gel bands were subjected to in-gel protein digestion with trypsin, following standard protocols. LC-MS/MS analyses were carried out using a nanoflow reversed phase HPLC (Agilent) and an LTQ ion trap mass spectrometer (Thermo Electron). Protein identifications were carried out by database search using Sequest software (Thermo Electron) against an *E. coli* protein database, generated

from the Uniref100 protein database. The protein sequence of the MBP-GS-GFP fusion protein was added to the *E. coli* protein database.

2.7.12 Identity of the Bands at ~67 and ~134 kDa

Proteins in the bands at ~67 and ~134 kDa were identified by in-gel digestion and LC-MS/MS analysis.

2.8 Acknowledgements

Norville, Derda, and Knight conceived of and developed the idea of using BioScaffolds for circuit optimization (i.e., creating the biological equivalent of electronic integrated circuits.) Norville, Derda, Belcher, Leschziner, and Knight developed the idea of using BioScaffolds to create protein fusions using BioBrick standard biological parts rather than new assembly standards. Norville, Belcher, and Knight conceived of and developed the initial BioScaffold standard. Norville, Belcher, and Leschziner conceived of and developed the idea of using visual screening to distinguish properly excised BioScaffolds from BioScaffolds that remained in the circuit. Norville, Gupta, and Knight conceived of and Norville and Gupta performed the flow cytometry experiments. Norville and Drinkwater conceived of and created the BioScaffold prototype, performed the assemblies, and excised and replaced the BioScaffold with RBSs. Norville, Derda, Gupta, Drinkwater, Belcher, Leschziner, and Knight drafted the manuscript.

Chapter 3

The S-Layer Protein SbpA for Modular Design of Protein Assemblies

In Chapter 1 on page 19, I described both engineered and modular components that can be brought together to create more complex systems. I also describe bacterial microcompartments, which are protein cages that bring the enzymes for metabolic pathways together in close proximity. The S-layer protein SbpA has properties that potentially allow it to perform functions analogous to bacterial microcompartments, yet are easier to engineer. A fundamental challenge in using SbpA to bring together enzymes in order to increase the efficiency of metabolic reactions is that recombinant SbpA is challenging to purify and that the purification process can decrease the efficiency of attached enzymes for *in vitro* reactions.

The S-layer protein SbpA contains a standard interface that allows the protein to self-assemble into crystals. In Section 1.2.2 on page 23 of Chapter 1, I discussed how the standard twistlock fittings and fixed sizes of shipping containers makes it easy to link the containers together (or to a transport chassis) on demand. Similarly, the S-layer protein SbpA crystallizes in 2D when a divalent cation is added to protein in solution; conversely, in the absence of divalent cations, SbpA remains monomeric in solution. A variety of proteins have been linked to the C-terminus of SbpA as fusion

proteins, yet do not interfere with the domains of the protein that self-assemble; thus, SbpA fusion protein can in many cases form an ordered array when calcium is added to solution, regardless of the identity of the fusion protein.

3.1 Structure and Content

In this chapter, I examine the structure of crystallized SbpA (Section 3.9 on page 91), the conformation of SbpA before crystallization (Section 3.2), as well as how to both purify soluble protein and perform a simple activity assay (Section 3.2).

This chapter contains two manuscripts. The first manuscript, "Fast and easy protocol for the purification of recombinant S-layer protein for synthetic biology applications." [265], introduces a non-denaturing purification protocol, a simple assay for crystallization, and the properties of SbpA before crystallization. The second manuscript "7 Å projection map of the S-layer protein SbpA obtained with trehalose-embedded monolayer crystals." [263] investigates the properties of SbpA during and after crystallization and introduces a new preparation method that yields the highest resolution lipid monolayer crystal to date.

3.2 A Soluble Variant of the S-Layer Protein SbpA for Modular Design: "Fast and Easy Protocol for the Purification of Recombinant S-Layer Protein for Synthetic Biology Applications"

Sections 3.3 to 3.8 contain material from the manuscript "Fast and easy protocol for the purification of recombinant S-layer protein for synthetic biology applications." written by Julie E. Norville, Deborah F. Kelly, Thomas F. Knight, Jr., Angela M. Belcher, and Thomas Walz that was published in the Biotechnology Journal in 2011. The Appendix A on page 143 contains additional information from the manuscript. The original abstract is included in the final dissertation summary in Section 5.1.2

on page 138. For reference, Section 1.5 on page 48 contains a list of abbreviations used throughout the dissertation.

3.3 Introduction

A major goal in synthetic biology is to simplify the design and construction of biological systems [283]. SbpA [87, 270, 359], the S-layer protein of *Lysinibacillus sphaericus* (*L. sphaericus*), and some other S-layer proteins are of particular interest to the synthetic biology community due to their modular structure and their ability to form 2D arrays in a divalent-cation-dependent manner [119]. Although a non-denaturing purification protocol exists for the S-layer protein RsaA from *Caulobacter crescentus* (*C. crescentus*) and RsaA fusion proteins [57], purification of both native SbpA from *L. sphaericus* and recombinant SbpA (expressed in *E. coli*) involves the use of denaturing conditions [57, 87, 119, 166, 257, 263, 270, 336], often followed by lyophilizing and refolding the denatured protein by dialysis.

Herein, we present a simple method a one-step, non-denaturing, nickel-affinity purification procedure that makes it possible to produce milligram amounts of a recombinant, His-tagged version of SbpA (SbpA-His_7; residues 31-1068 [257]) as well as other N- and C-terminally truncated variants in *E. coli*. Previously, SbpA has been expressed at 37°C (see Supporting Information in Appendix A on page 143). Here, expression in *E. coli* at low temperature (see Supporting Information in Appendix A on page 143) allowed SbpA to fold properly and made it possible, for the first time, to purify recombinant SbpA under non-denaturing conditions. The recombinant SbpA variants retain their ability to form crystalline arrays in solution, which in the examples examined correlate with the appearance of a white precipitate.

3.4 Materials and Methods

3.4.1 Preparation of Recombinant SbpA Protein under Non-Denaturing Conditions

The gene encoding SbpA (residues 31-1268) was amplified by PCR from genomic DNA of *L. sphaericus* (ATCC number 4245) and ligated into the NdeI and BamHI restriction sites of the pET28a vector (EMD Biosciences) with a C-terminal 7 x His tag. From this construct, the three constructs used in this study were generated: SbpA-His_7-encoding residues 31-1068, which is a construct commonly used in engineering applications [169, 191, 257, 332]; C-terminal truncation mutant $SbpA_{C-trunc}$-His_7, encoding residues 31-918 [71]; and N-terminal truncation mutant $SbpA_{N-trunc}$- His_7, encoding residues 203-1068 [71].

Each construct was expressed in *E. coli* HMS174 (DE3) (EMD Biosciences). The culture was grown at 37°C in LB medium containing 50 μg/mL kanamycin and 50 μg/mL rifampicin until the OD_{600} reached 0.6. Upon induction, by addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), the incubation temperature was lowered to 18°C. 18 h after induction, cells were pelleted by centrifugation at 16,000 g for 15 min at 4°C. The cell pellet was resuspended in lysis buffer [129] containing 20 mM Tris, pH 8.0, 300 mM NaCl, 10 mM imidazole, and 1 mg/mL lysozyme (EMD Biosciences) and sonicated on ice. Insoluble debris was removed by centrifugation at 20,000 g for 30 min at 4°C. The cleared lysate was incubated with nickel nitrilotriacetic acid (Ni-NTA) agarose beads (Qiagen, Valencia, CA) for 1 h at 4°C or centrifuged through a Ni-NTA spin column (Qiagen, Valencia, CA). Following washing with buffer containing 100 mM imidazole in 20 mM Tris, pH 8.0, 300 mM NaCl, SbpA was eluted in buffer containing 250 mM imidazole as described [129]. The eluted protein was finally dialyzed at 4°C against 20 mM Tris, pH 8.0.

3.4.2 Purification of Native SbpA

Native SbpA was prepared as described previously [39, 263].

3.4.3 SDS-PAGE, Western Blotting, Quantitation

SDS-PAGE

Samples were run on 10% Bis-Tris gels (Invitrogen) and stained with Coomassie blue stain.

Western Blots

His-tagged proteins were detected with anti-His antibody (GE Healthcare) and blots were developed with the alkaline phosphatase method by using the Sigma Fast system (Sigma-Aldrich) [189].

Quantitation

The ratios of SbpA-His_7 in inclusion bodies and the soluble fraction were analyzed by using quantitatively stained SDS-PAGE gels [68].

3.4.4 Specimen Preparation, Electron Microscopy and Image Processing

Negative Staining

Purified SbpA and 2D crystals of SbpA formed in solution or on lipid monolayers were stained with uranyl formate by using the conventional negative-staining technique described in reference [267].

Electron Microscopy (EM)

Samples were imaged with an FEI Tecnai 12 electron microscope (FEI, Hillsboro, OR) equipped with an LaB_6 filament and operated at an acceleration voltage of 120 kV. Images were recorded on imaging plates under low-dose conditions at a nominal magnification of 67,000x and a defocus value of about 1.5 μm. Imaging plates were scanned with a Ditabis scanner (Pforzheim, Germany) at a step size of 15 μm, a

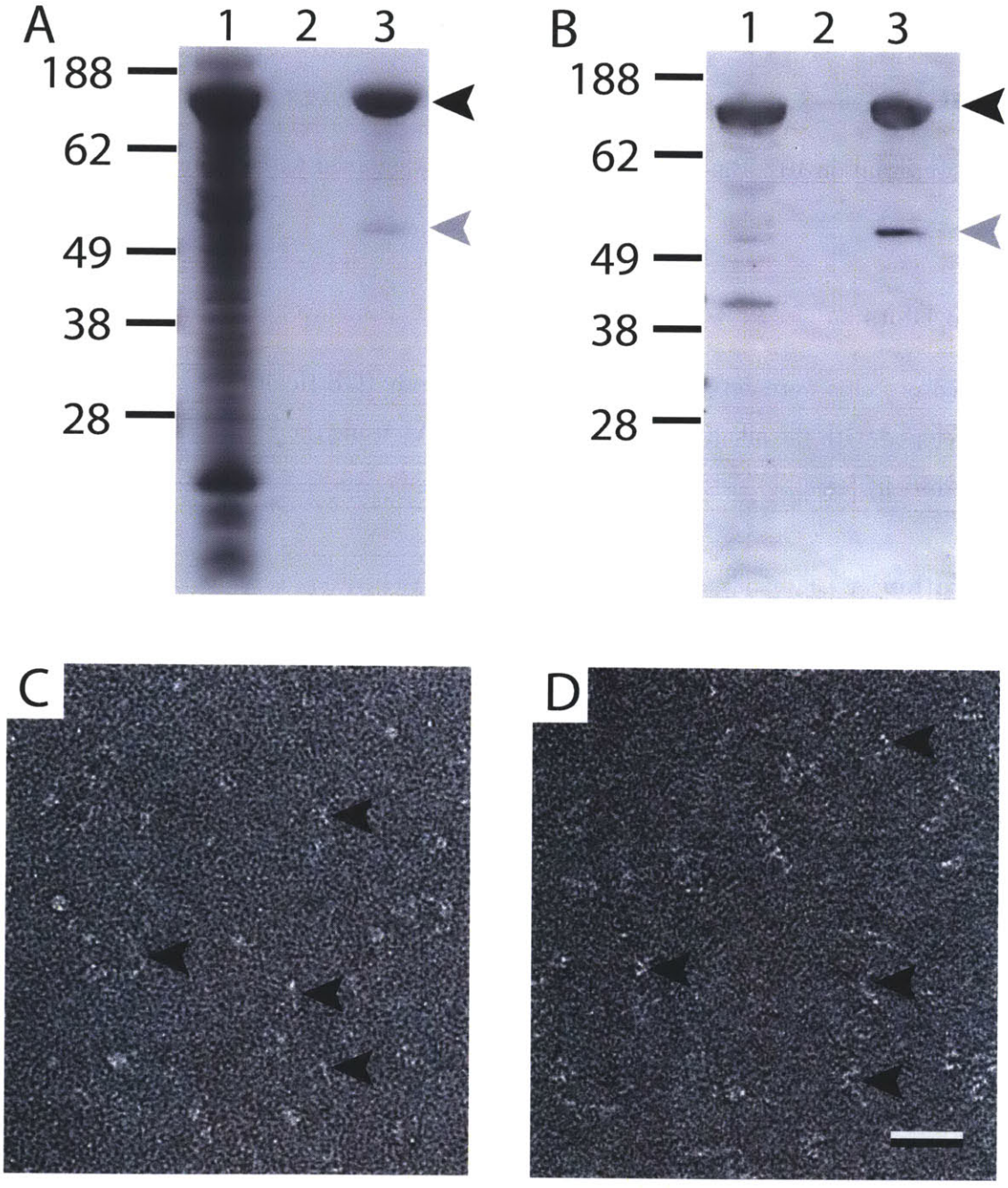


Figure 3-1: Expression and Purification of Recombinant SbpA-His$_7$ Containing Residues 31 to 1068. The non-denaturing purification process was followed by *(A)* SDS-PAGE and *(B)* western blot analysis with anti-His antibody. Lane 1: cell lysate, Lane 2: 100 mM imidazole wash, Lane 3: 250 mM imidazole elution. *(C)* Negative-stain EM image of SbpA-His$_7$ purified by using the non-denaturing protocol. *(D)* Negative-stain EM image of full-length native SbpA purified by using a standard, denaturing protocol. The black arrows in *(C)* and *(D)* indicate protein in the extended conformation, which stains poorly and is difficult to see. Scale bar: 50 nm. Experiments were reproducible within ten different preparations.

gain setting of 20,000, and a laser power setting of 30% [222]. The images were then binned over 2 x 2 pixels for a final sampling of 4.5 Å/pixel at the specimen level.

Image Processing

Fourier transforms were calculated using the 2dx software package [140].

3.5 Results and Discussion

To produce SbpA under non-denaturing conditions, we amplified the SbpA gene from the genomic DNA of *L. sphaericus* (ATCC number 4525) [257, 336], cloned it into the pET28a vector by following an established protocol [257], and expressed the protein in *E. coli* strain HMS174(DE3) (EMD Biosciences). Once induced, cells were cultured for 18 h at 18°C, and then harvested and lysed as described before [129].

The cell lysate containing the recombinant SbpA-His_7 protein (Figures 3-1 on the preceding page A and B, lane 1) was incubated with Ni-NTA affinity resin for 1 h at 4°C. After washing with buffer containing 100 mM imidazole (Figures 3-1 on the facing page A and B, lane 2), the protein was eluted with the same buffer containing 250 mM imidazole (Figures 3-1 on the preceding page A and B, lane 3). SDS-PAGE analysis showed that the one-step purification procedure yielded recombinant SbpA, running at an apparent molecular weight of 127 kDa (Figure 3-1 on the facing page A, lane 3, black arrow), that was over 95% pure. A faint band at 53 kDa seen on the gel (Figure 3-1 on the preceding page A, lane 3, gray arrow) was also detected by the anti-His antibody (Figure 3-1 on the facing page B, lane 3), and therefore, must have represented a C-terminal fragment of SbpA-His_7. For a comparison of the yield of different SbpA variants and a description of how SbpA-His_7 partitioned between the soluble fraction and inclusion bodies at different expression temperatures, see Supporting Information (i.e., Section A.1.1 on page 143), Figures A-1 on page 145 and A-1 on page 145.

To characterize recombinant SbpA-His_7, we visualized the purified protein by negative-stain EM (see Supporting Information (Section A.1 on page 143) and the

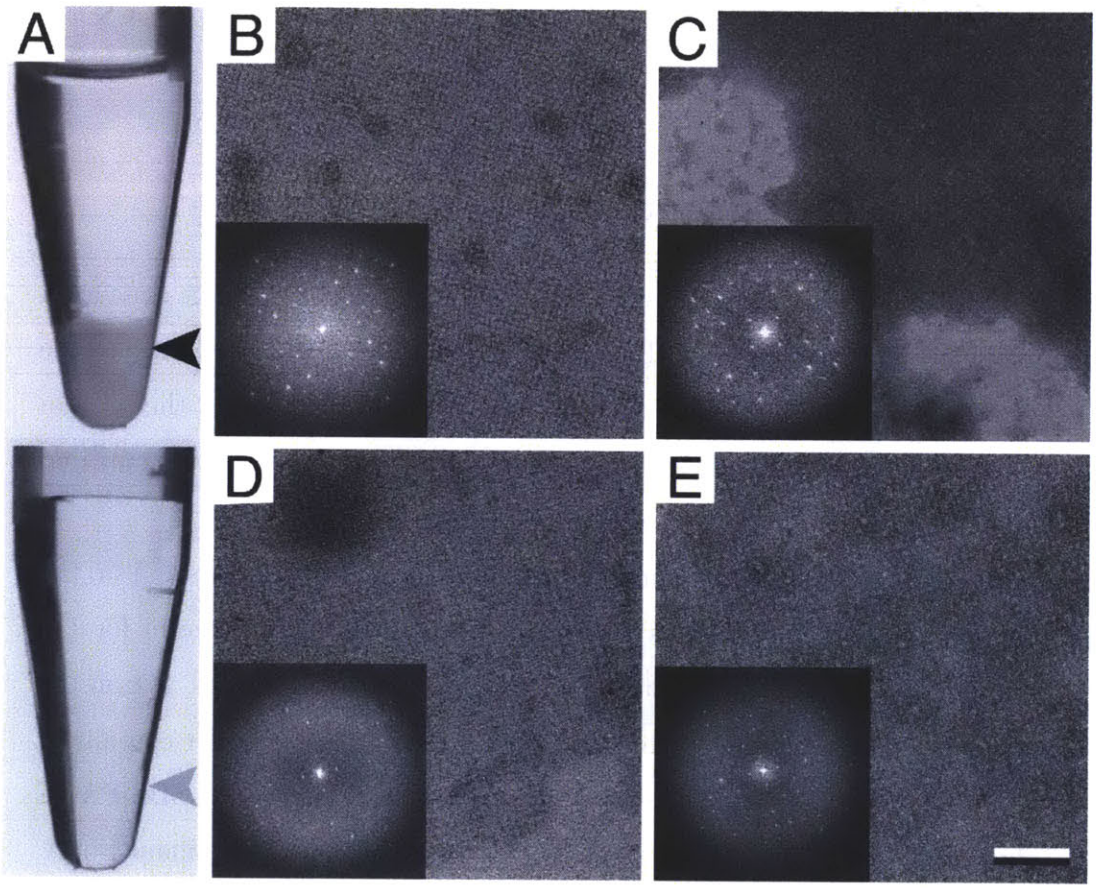


Figure 3-2: Native and Recombinant SbpA are Active and Form 2D Crystals in Solution. *(A)* Addition of 50 mM $CaCl_2$ to full-length native SbpA results in a white precipitate (top panel) that does not occur in the absence of $CaCl_2$ (bottom panel). Images of crystalline arrays formed by SbpA variants in solution: *(B)* SbpA-His_7 (residues 31-1068), *(C)* $SbpA_{C-trunc}$-His_7 (residues 31-918), *(D)* $SbpA_{N-trunc}$-His_7 (residues 203-1068), and *(E)* native SbpA. Scale bar: 100 nm. SbpA-His_7 and native SbpA experiments were reproducible within ten different preparations. $SbpA_{C-trunc} - His_7$ and $SbpA_{N-trunc}$-His_7 were reproducible within three different preparations.

Materials and Methods (Section 3.4 on page 82) for details). Although the protein stained poorly and was thus difficult to see, the images revealed a heterogeneous particle population with particles adopting either a compact or a flexible, extended conformation (Figure 3-1 on page 84 C). To determine whether the observed heterogeneity was an inherent characteristic of SbpA or specific to the recombinant protein, we also purified native SbpA from the cell walls of *L. sphaericus* (ATCC 4525) by using a denaturing procedure [263]. Images of negatively stained samples of native SbpA showed a similarly heterogeneous particle population (Figure 3-1 on page 84 D).

To assess the activity of both native and recombinant SbpA, we tested the two proteins for their ability to form 2D crystals in solution (see Supporting Information (Section A.1 on page 143) and the Materials and Methods (Section 3.4 on page 82) for details). Addition of $CaCl_2$ to SbpA induces the protein to crystallize in solution, promoting the formation of self-assembly products [71, 257, 288]. When we added 50 mM $CaCl_2$ to native SbpA at concentrations of 12 mg/mL, we observed the formation of a white precipitate that settled over time at the bottom of the sample tube (Figure 3-2 on the facing page A, top panel). For both native and recombinant SbpA at 7.6 nmol (i.e., 1 mg/mL for native SbpA), the white precipitate appeared after 18 h, and no differences could be detected in the rate at which the precipitates formed and in the final amounts of precipitates. Since precipitates never occurred in the absence of $CaCl_2$ (Figure 3-2 on the preceding page A, bottom panel), we speculated that the white precipitate was caused by the formation of SbpA crystals. This notion was verified by negative-stain EM, which only showed 2D crystals in SbpA samples that contained the white precipitate. The same white precipitate also formed when 50 mM $CaCl_2$ was added to solutions containing different recombinant His-tagged SbpA versions, namely, a long construct often used in bio-engineering, SbpA-His_7 (residues 31-1068), an N-terminal truncation mutant, $SbpA_{N-trunc}-His_7$ (residues 203-1068), and a C-terminal truncation mutant, $SbpA_{C-trunc}$-His_7 (residues 31-918), which were all expressed in *E. coli* and prepared using our non-denaturing purification protocol.

To assess whether the recombinant proteins purified under non-denaturing condi-

tions were functional, we used negative-stain EM (see Supporting Information (Section A.1 on page 143) and Materials and Methods (Section 3.4 on page 82) for details) to compare the quality of 2D arrays formed by the recombinant proteins in solution with that of arrays formed by native SbpA (purified by using a denaturing protocol). We prepared three samples each for native SbpA and the recombinant proteins (crystallized at a 7.6 nmol concentration) and found that 2D crystals were prevalent on all of the grids (typically 1014 crystals were present per grid square). The appearances of crystals formed by SbpA-His_7 (Figure 3-2 on page 86 B), $SbpA_{C-trunc}$-His_7 (Figure 3-2 on page 86 C), $SbpA_{N-trunc}$-His_7 (Figure 3-2 on page 86 D), and native SbpA (Figure 3-2 on page 86 E) were also indistinguishable. All arrays had side lengths that ranged from 0.5 to 2 μm, but the arrays formed by $SbpA_{N-trunc}$-His_7 were on average 30% smaller than those formed by the other SbpA variants. Furthermore, in lipid monolayer crystallization experiments, N-terminally truncated $SbpA_{N-trunc}$-His_7 consistently failed to form crystals on lipid monolayers (see the Supporting Information, Figure A-3 on page 147). These results suggest that the N-terminus of SbpA is predominantly required for surface binding, but may also play a minor role in promoting array formation.

Fourier transforms calculated for five images of each sample of 2D crystals formed in solution indicated that all crystals had a similar degree of order with diffraction spots visible to a resolution of at least 25 Å (insets in Figures 3-2 on page 86 BE), which was close to the resolution limit imposed by the negative stain. The lattice constants for arrays formed by the various SbpA proteins were a = b = (131 ± 0.4) Å for SbpA-His_7, a = b = (131 ± 3) Å for $SbpA_{N-trunc}$-His_7, and a = b = (129 ± 4) Å for $SbpA_{C-trunc}$-His_7 compared with a = b = (131 ± 0.9) Å for the native protein.

Examination of purified native SbpA and recombinant SbpA-His_7 by negative-stain EM revealed that the particle population in both preparations was heterogeneous. Although more of the SbpA-His_7 particles appeared to be in a compact conformation (Figure 3-1 on page 84 C) and more of the native SbpA particles appeared to be in an extended, flexible conformation (Figure 3-1 on page 84 D), both protein preparations were equally effective in forming 2D arrays in solution (Figure 3-2 on

page 86). Thus, the different appearance of the two protein preparations does not appear to represent proteins with different properties, but may simply indicate that SbpA is in a dynamic equilibrium that can shift between a compact and extended, flexible conformation.

All SbpA proteins tested formed crystals in solution and the presence of SbpA crystals in solution correlated, in each case, with the formation of a white precipitate in the sample solution. Previously, however, an expressed C-terminal truncation mutant of SbpA was purified by using a denaturing purification procedure and formed crystals that displayed *p1* rather than *p4* symmetry [71]. In contrast, crystals formed by our identical $SbpA_{C-trunc}$-His_7 construct (except for an additional histidine residue in the affinity tag) that were isolated by using our non-denaturing purification procedure still formed crystals with native *p4* symmetry (Figure 3-2 on page 86 E). This result suggests that the loss of fourfold symmetry observed in the previous study does not result from C-terminal truncation, but from the harsh purification conditions used to purify the truncation mutant.

3.6 Concluding Remarks

Our non-denaturing purification procedure to prepare recombinant active SbpA variants in milligram amounts is likely to aid in the engineering and production of new biomaterials. In the field of synthetic biology, most chemical biosynthesis reactions are currently performed by using enzymes produced within cells, which can be problematic if toxic products are produced [251], if the ratio of different enzymes needs to be tuned genetically [264], or if enzymes have to be brought together to obtain high yields [109]. Our protocol may be particularly advantageous to produce SbpA with C-terminally fused enzymes, which are likely to be more challenging to refold. Such fusion proteins would make it possible to bring different enzymes together by assembling them onto beads [74] or in solution and to vary their ratios simply by altering the concentration of each SbpA fusion protein before assembly. The ability to purify SbpA under non-denaturing conditions thus makes SbpA a powerful tool

for *in vitro* synthetic biology.

3.7 Contributions

Norville, Belcher, and Knight conceived of using SbpA as a model biomaterial biological part for synthetic biology. Norville, Kelly, Belcher, Knight, and Walz designed the experiments. Norville, Kelly, and Walz wrote the manuscript.

3.8 Acknowledgements

This work was supported by a KAUST Scholar Graduate Research Fellowship (J. E. N.), the SynBERC NSF ERC (www.synberc.org) (J. E. N. and T. F. K.), National Institutes of Health grant GM52580 (S. C. Harrison) and the U.S. Army Research Office through both the Institute for Soldier Nanotechnologies and the Institute for Collaborative Biotechnologies (A. M. B.). T. W. is an investigator in the Howard Hughes Medical Institute.

3.9 Examination of Crystalline SbpA at the Highest Resolution to Date: "7 Å Projection Map of the S-Layer Protein SbpA Obtained with Trehalose-Embedded Monolayer Crystals"

Sections 3.10 to 3.15 on page 112 contain material from the manuscript "7 Å projection map of the S-layer protein SbpA obtained with trehalose-embedded monolayer crystals" written by Julie E. Norville, Deborah F. Kelly, Thomas F. Knight, Jr., Angela M. Belcher, and Thomas Walz that was published in the Biotechnology Journal in 2011. The original abstract is included in the final dissertation summary in Section 5.1.3 on page 139. For reference, Section 1.5 on page 48 contains a list of abbreviations used throughout the dissertation.

3.10 Introduction

Fromherz was the first to demonstrate that soluble proteins can form ordered arrays on the surface of a lipid monolayer at the air/water interface and that these arrays can be imaged by electron microscopy (EM) [128]. Subsequently, the lipid monolayer crystallization method was used to make soluble proteins amenable to structural studies by electron crystallography [341]. The underlying principle of array formation is that association of a target protein with the lipid monolayer leads to concentration and partial alignment of the protein. Since the lipid monolayer is in its fluid phase, the lipids, and hence the associated proteins, can diffuse in the plane of the monolayer, allowing the proteins to interact with each other and under favorable conditions to form regular arrays. The lipid monolayer is usually formed with neutral lipids spiked either with charged lipids to induce association of proteins by electrostatic interactions [100, 249, 334, 335] or with functionalized lipids, such as lipids containing Ni-NTA headgroups that specifically recruit His-tagged proteins to the monolayer [204]. Once introduced, the method was subsequently adapted for use with membrane proteins

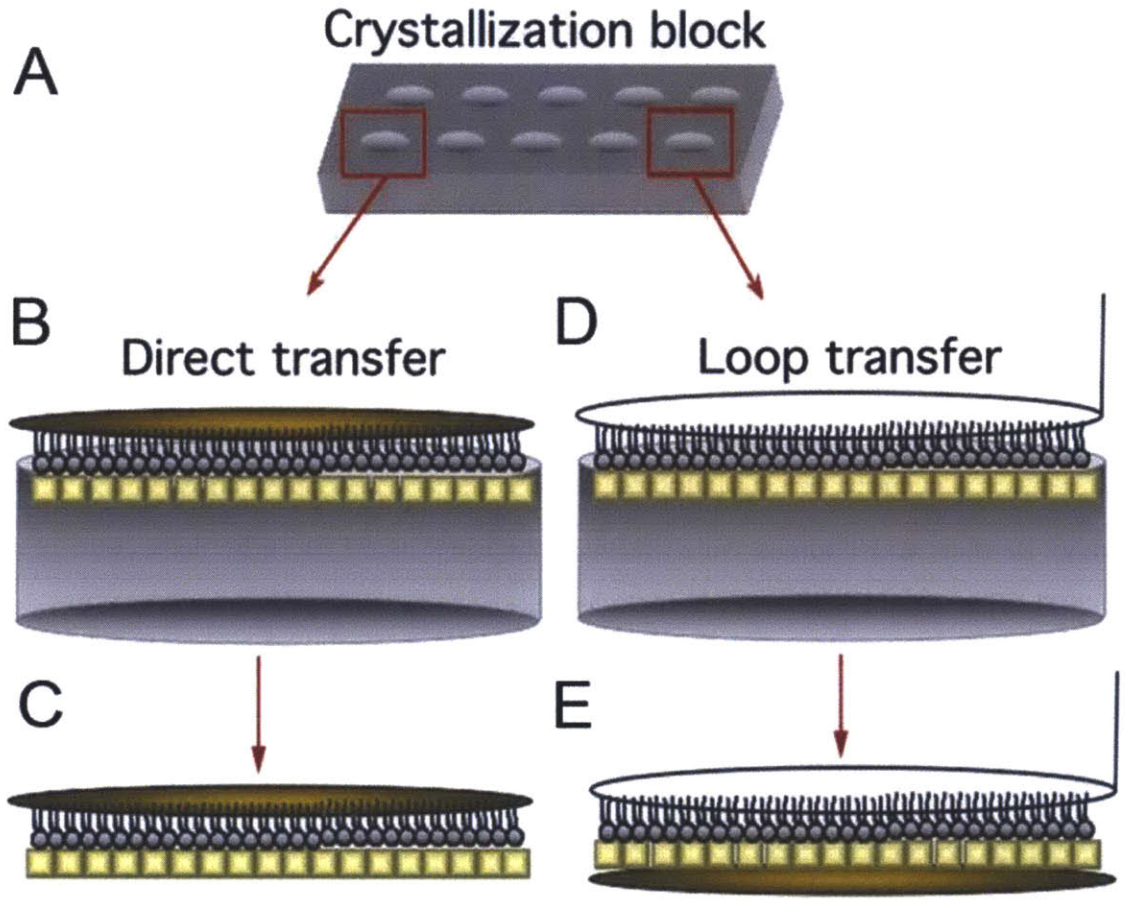


Figure 3-3: The Direct and Loop Transfer Methods. *(A)* Teflon block used to grow 2D crystals on lipid monolayers. *(B, C)* Direct transfer method, after which the lipid monolayer is in contact with the carbon film. *(D, E)* Loop transfer method, after which the protein crystal is in contact with the carbon film. For *(B E)* Protein crystals are in yellow while the EM grid is colored gold.

[205, 216] and for the assembly of protein complexes [81, 187, 188].

Although it was possible to record high-resolution (~3 Å) EM data of streptavidin crystals on a lipid monolayer [37, 203], a persistent problem in the use of this technique lies in the transfer of lipid monolayer crystals to an EM grid without deteriorating the order of the protein arrays. In addition, lipid monolayer samples are often not flat, and to date no high-resolution images have been reported for tilted lipid monolayer crystals. A reliable protocol to reproducibly prepare specimens for high-resolution EM data collection would thus have the potential to boost interest in lipid monolayer

crystallization for structural studies of soluble proteins.

Once crystals have formed on the lipid monolayer, often using a Teflon crystallization block (Figure 3-3 on the facing page A), they can be transferred to an EM grid using the direct transfer method (Figures 3-3 on the preceding page B and C), historically referred to as the Langmuir-Schaefer transfer method [207]. For this technique, a hydrophobic carbon film mounted on an EM grid is brought into direct contact with the hydrophobic tails of the lipids in the monolayer (Figure 3-3 on the facing page B). When the grid is lifted off, the monolayer and hence the associated crystals adhere to the carbon film and are thus transferred to the grid (Figure 3-3 on the preceding page C). The crystals on the grid can then be blotted and either negatively stained or vitrified for subsequent observation in the electron microscope. However, touching a lipid monolayer with a carbon film imparts mechanical force that tends to distort the attached crystals [67]. Thus, a holey grid is often used in conjunction with direct transfer and specimen vitrification. Images are then taken in regions without carbon film, where the crystals should be least affected. Although variations of this technique have yielded both the highest resolution lipid monolayer crystals to date [37, 203] as well as 3D reconstructions at lower resolutions [188, 223–225, 331, 357], high-resolution 3D structures have proved difficult to obtain in this manner. An alternative approach to the direct transfer method is the loop transfer technique, in which a wire loop is used to lift off the lipid monolayer and to deposit it on an EM grid (Figures 3-3 on the facing page D and E). This method was initially developed with holey carbon grids [36], but a continuous carbon support film can also be used. In contrast to the direct method, when loop transfer is used, the protein array rather than the lipid monolayer makes contact with the carbon film (Figure 3-3 on the preceding pageE), leading to a more robust interaction between the protein crystal and the carbon support film.

In a related field, electron crystallography of conventional 2D crystals, i.e., membrane proteins reconstituted into lipid bilayers, a number of high-resolution structures have now been determined [142, 155, 159, 160, 193, 252, 346]. Advances in specimen preparation, in particular the embedding of the specimen in sugar solutions

[98, 147, 158, 172, 340], have played a crucial role in the success of this method. It was thus of interest to us, whether sugar embedding could be adapted to the preparation of lipid monolayer crystals and how it would compare to the conventional preparation methods currently used for lipid monolayer crystals. As a test specimen to study specimen preparation methods, we selected the S-layer protein, SbpA. S-layer proteins have been historically used for electron crystallographic studies due to their inherent propensity to self-assemble into ordered arrays (for a review, see [48]). In particular, we chose the protein SbpA from *L. sphaericus*, because it is easy to purify from the cell wall of native bacteria [39] and it has been shown to form 2D crystals [286] in a calcium-dependent manner [287]. Native and recombinant S-layer proteins have also recently been used as building blocks to design 2D scaffolds with engineered properties for various nanotechnology applications [320], raising interest in determining their structure. However, structural studies of SbpA crystals have so far not exceeded a resolution of about 20 Å [212, 213]. By systematically testing various specimen preparation protocols, we found that we could produce the best specimens by transferring the lipid monolayer crystals to a continuous carbon film using loop transfer followed by embedding the sample in trehalose and quick-freezing in liquid ethane. With images of specimens prepared in this way, we could calculate a projection map of SbpA at a resolution of 7 Å. This is one of the highest resolution maps obtained with 2D crystals grown on lipid monolayers to date. At this point the resolution may no longer be limited by the specimen preparation technique but by the size and order of the SbpA arrays themselves.

3.11 Materials and Methods

3.11.1 Purification of SbpA

SbpA was purified as described [39] with minor modifications. Briefly, *L. sphaericus* (ATCC No. 4525) was grown at 32°C in SVIII medium (50 mM Hepes, pH 7.2, 7 mM K_2HPO_4, 10 g/l peptone, 5 g/l yeast extract, 5 g/l meat extract, 0.2 mM

$MgSO_4$, 1.8 mM sucrose, 17 mM glucose). The cells were lysed by sonication in 50 mM Tris, pH 7.2, and the cell walls were isolated by centrifugation at 16,000g for 15 min. The pellet was resuspended in Buffer A (0.75% Triton X-100 in 50 mM Tris, pH 7.2) with a Tissue Tearor (BioSpec Products, Bartlesville, OK) and centrifuged at 28,000g for 10 min. The cell walls were washed three more times in Buffer A. SbpA (1.5-2.0 mg/ml) was released from the cell walls by unfolding the protein in Buffer B (50 mM Tris, pH 7.2, 5 M guanidine HCl) for 30 min at room temperature with stirring. The cell walls were removed by centrifugation at 100,000g for 45 min. SbpA in the supernatant was refolded by dialysis against distilled water and aggregates were removed by centrifugation at 100,000g for 30 min.

3.11.2 Lipid Monolayer Crystallization of SbpA

1,2-Dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) was purchased from Avanti Polar lipids (Alabaster, AL) and didodecyldimethylammonium bromide (DDMA) was purchased from Acros Organics (Geel, Belgium). To set up monolayer crystallization trials Teflon blocks with 10-18 wells were used, each well being 5 mm in diameter and not, ~1 mm in depth. The total volume of sample placed into each well was ~25 µl (Figure 3-3 on page 92A). For crystallization, 23.1 µl of crystallization buffer containing 20 mM Tris, pH 8, 100 mM NaCl and 50 mM $CaCl_2$ was added to the wells. Lipid solution was prepared by mixing 20% DDMA (1 mg/ml) in chloroform with 55% DMPC (1 mg/ml) in chloroform with 25% chloroform. One microliter of this lipid mixture was added on top of the protein solution in each well. After approximately 15 min of incubation at room temperature, a Hamilton syringe was used to inject 1.9 µl of SbpA protein in deionized water (1 mg/ml) into the aqueous sub-phase to initiate crystallization. The crystallization blocks were incubated in a sealed humid environment at room temperature. While crystals formed within 2 h of incubation and were stable for at least 2 months, samples for cryo-EM experiments were incubated for 1 week before specimens were prepared.

3.11.3 Lipid Monolayer Transfer Techniques

Lipid monolayer crystals were transferred to continuous carbon film mounted on copper EM grids (400 mesh, Ted Pella, Redding, CA) or to holey carbon grids (400 mesh, 2/4 Quantifoil Micro Tools GmbH, Germany) using both the direct transfer [341] and loop transfer [36] methods. For direct transfer, an EM grid (without prior glow discharging) was carefully placed on top of the lipid monolayer sample in the crystallization block and gently picked up again with a pair of forceps. The grid was blotted with Whatman #1 filter paper (Whatman International Ltd., Middlesex, England) and either negatively stained with 1% uranyl acetate or plunge-frozen in either liquid nitrogen or liquid ethane. For the loop transfer, a thin platinum loop with a diameter of 3.5 mm (Ernest F. Fullam Inc., Latham, NY) was brought gently into contact with the lipid monolayer sample in the crystallization block. The loop was then lifted up and the sample applied to a glow-discharged carbon film mounted on an EM grid. The grid was blotted and negatively stained or plunge-frozen as described above.

3.11.4 Sample Preparation

We used six methods to prepare frozen-hydrated specimens of lipid monolayer SbpA 2D crystals. Direct transfer of 2D crystals to (1) holey and (2) continuous carbon grids, followed by blotting for 10 s and plunge-freezing into liquid ethane; loop transfer to (3) holey and (4) continuous carbon grids, followed by blotting for 10 s and plunge-freezing into liquid ethane; loop transfer to continuous carbon grids, followed by embedding with 5% trehalose solution and plunge-freezing either into (5) liquid nitrogen or (6) liquid ethane.

3.11.5 Data Collection

Negatively stained specimens were imaged in a T12 electron microscope (FEI, Hillsboro, OR) equipped with a LaB_6 filament and operated at an acceleration voltage of 120 kV. Images were taken on a Gatan Ultrascan 894 2Kx2K charge-coupled device

(CCD) camera. Grids of frozen-hydrated SbpA crystals were transferred into a F20 electron microscope (FEI, Hillsboro, OR) equipped with a field emission gun using an Oxford cryo-specimen holder, maintaining a temperature of ~180°C. Samples were examined at an acceleration voltage of 200 kV and images of crystals were recorded at a magnification of 50,000x using low-dose procedures. Images were recorded either with a Gatan MegaScan 794 2Kx2K CCD camera (for general characterization of the samples) or on Kodak SO-163 film (for quantitative analyses). Films were developed for 12 min with full-strength Kodak D-19 developer at 20°C.

3.11.6 Image Processing

An optical diffractometer was used to select for drift-free images and to identify the best diffracting regions of the imaged SbpA 2D crystals. Micrograph areas (5000 x 5000 pixels) of SbpA 2D crystals were digitized with a Zeiss SCAI scanner (Carl Zeiss Inc., Oberkochen, Germany) using a step size of 7-μm and processed with the 2dx software package [140], which included lattice unbending and correction for the contrast transfer function (CTF). The five best images, ranging in defocus from -0.8 to -2 μm, were merged and used to calculate a projection map with imposed *p4* symmetry using the MRC [96] and CCP4 [284] software packages. A negative temperature factor of -200 was applied to the projection map to enhance the high-resolution Fourier terms.

3.12 Results

3.12.1 2D Crystallization of SbpA

Ca^{2+} has been shown to induce SbpA to form regular arrays [287]. We could reproduce this array formation with SbpA purified from the cell wall of *L. sphaericus* (Figure 3-4 on the next page A). When protein was added to Ca^{2+}-free crystallization buffer, no SbpA arrays formed in the sample (data not shown), but when protein was added to Ca^{2+}-containing crystallization buffer small SbpA arrays appeared within 2 h (Figure

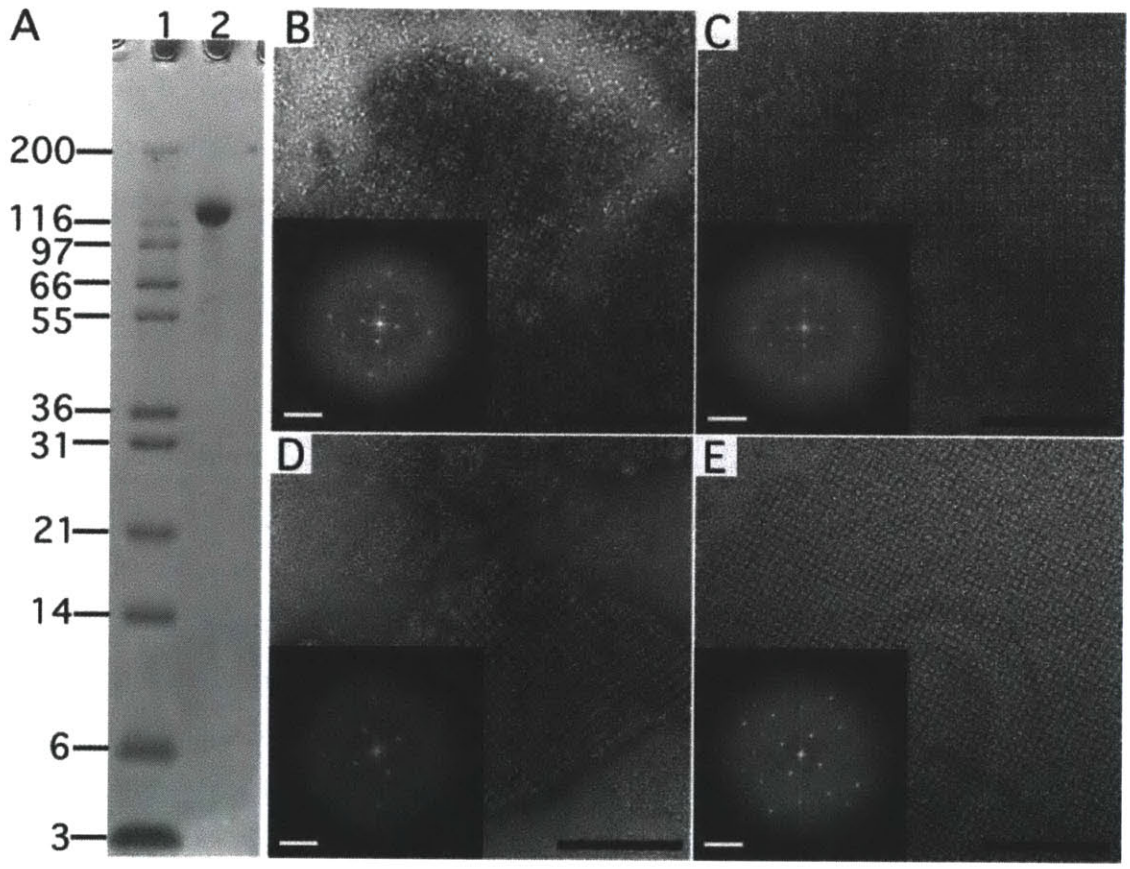


Figure 3-4: Negatively Stained SbpA Crystals Formed in Solution and on a Lipid Monolayer. *(A)* SDS PAGE gradient gel showing purified SbpA used for crystallization. Lane 1: Mark 12 protein standard (Invitrogen Corporation, Carlsbad, CA); lane 2: SbpA (MW ~120 kDa). *(B)* SbpA crystals formed in solution within 2 h after addition of 50 mM $CaCl_2$. *(C)* After incubation for 24 h, the SbpA crystals grew larger but the order of the SbpA arrays did not improve significantly (compare Fourier transforms shown as insets in panels *B* and *C*). *(D)* SbpA crystals also formed within 2 h of incubation with $CaCl_2$ on lipid monolayers. *(E)* After a 24 h incubation, the crystals grew not only bigger but also improved in order (compare Fourier transforms shown as insets in panels *D* and *E*). Scale bars in the images are 100 nm; scale bars in the Fourier transform are $(6.5\ \text{nm})^1$.

3-4 on the facing page B). While these arrays grew in size over time (Figure 3-4 on the preceding page C), the crystal order did not improve with time (compare Fourier transforms shown as insets in Figures 3-4 on the facing page B and C). Since SbpA has been shown to form regular arrays on positively charged lipid monolayers [286], we investigated whether the presence of a monolayer would induce the formation of better ordered 2D crystals. As before, initial SbpA arrays formed on lipid monolayers within two h of Ca^{2+} addition (Figure 3-4 on the preceding page D). However, in contrast to SbpA arrays formed in solution, the 2D crystals formed on the lipid monolayer not only grew in size over time (Figure 3-4 on the facing page E) but also showed a marked improvement in crystal order (compare Fourier transforms shown as insets in Figure 3-4 on the preceding page D and E). These results showed that 2D crystals formed on monolayers are of better quality and thus better suited for structural studies of SbpA than 2D crystals grown in solution.

3.12.2 Direct and Loop Transfer of SbpA Crystals onto EM Grids

We used negative staining to assess specimens prepared by the direct and loop transfer methods. For these experiments grids covered with a continuous carbon film were used. Representative images of crystals prepared with 1% uranyl acetate recorded at low magnification (8700x) show that large crystalline arrays of SbpA were present on the grid after direct transfer (Figure 3-5 on the following page A) as well as after loop transfer (Figure 3-5 on the next page B). Upon examining several grids prepared by the two transfer methods, we found that independent of the transfer method typically ~80% of the grid squares contained crystals. Grids prepared using the direct transfer method had, however, a greater number of smaller, broken crystals. A typical grid square contained between 5 to 8 crystals, which varied in size from 0.5 to 2 μm. This transfer method also tended to produce specimens, in which crystals were stacked on top of each other, resulting in several overlaid diffraction patterns in Fourier transforms of the images (data not shown). Grids prepared using the

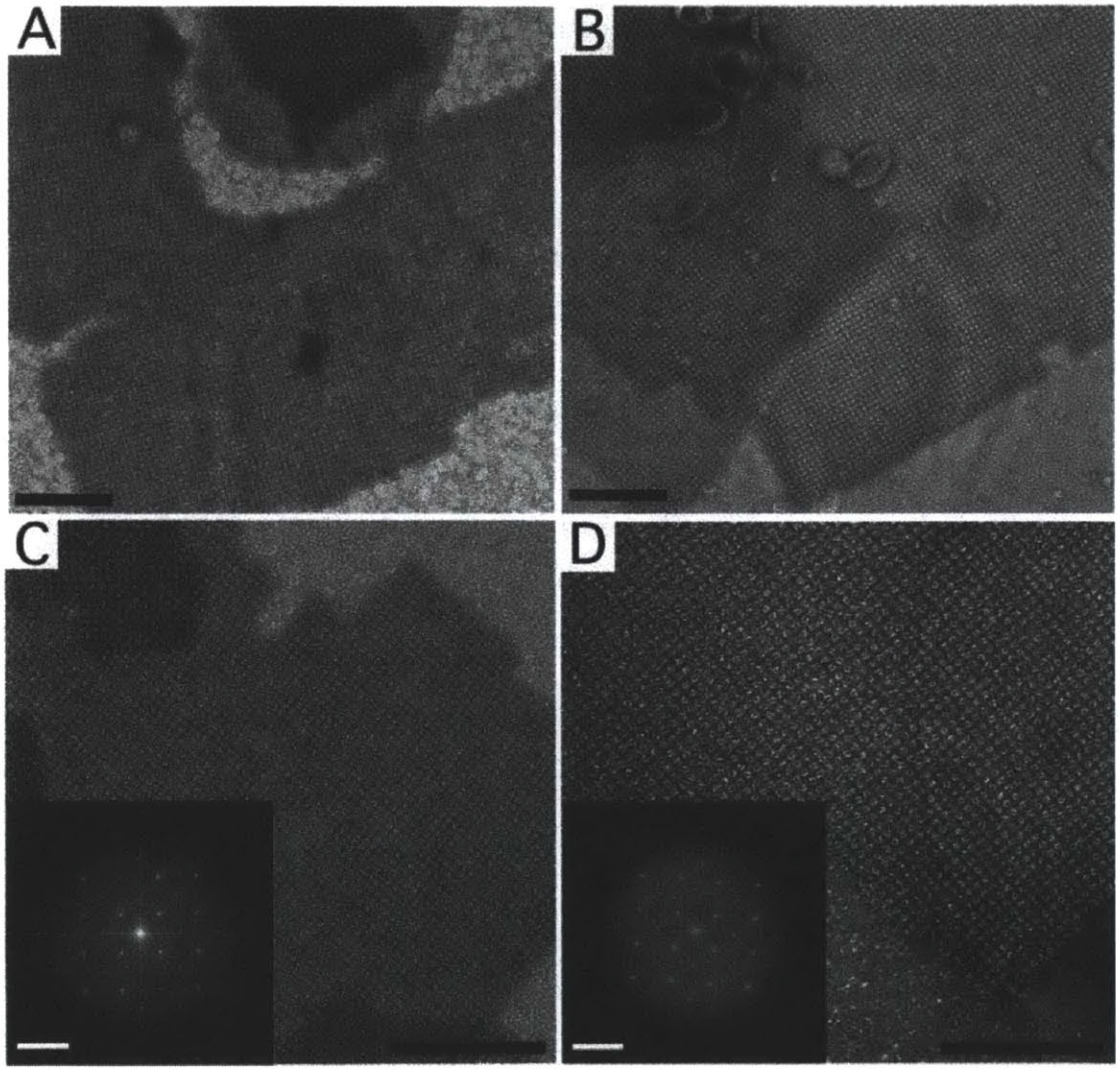


Figure 3-5: Negatively Stained SbpA Monolayer Crystals Prepared with the Direct and Loop Transfer Methods. *(A)* Low magnification image (8,700x) and (C) high magnification image (52,000x) of SbpA monolayer crystals prepared with the direct transfer method. *(B)* Low magnification image (8,700x) and *(D)* high magnification image (52,000x) of SbpA monolayer crystals prepared using the loop transfer method. The insets in panels *C* and *D* show the Fourier transforms of the crystals. Scale bars in the images are 100 nm; scale bars in the Fourier transform are $(6.5\text{ nm})^1$.

loop transfer showed about the same level of crystal coverage, but the crystals were in general less broken, as expected with this more gentle transfer technique. We typically found 2-4 crystals per grid square, which ranged in size from 3 to 5 μm. At higher magnification (52,000x) the square lattice formed by SbpA is clearly seen in negatively stained samples prepared by either method (Figures 3-5 on the facing page C and D), and Fourier transforms (insets in Figures 3-5 on the preceding page C and D) showed no difference in crystal order (number of diffraction orders in Fourier transforms) to the resolution limit introduced by the stain (~20 Å). Our results thus suggest that the direct and loop transfer methods are equally efficient in adsorbing SbpA 2D crystals to continuous carbon grids, without loss of crystal order to a resolution of at least 20 Å. The loop method appears, however, to induce less mechanical stress on the crystals during the transfer as indicated by the larger size of the observed crystals.

3.12.3 Image Processing

3.12.4 Frozen-Hydrated SbpA Crystals on Holey Carbon Film

In an attempt to obtain higher resolution information, we prepared frozen-hydrated specimens of the SbpA crystals using holey carbon film according to the protocol introduced by Kubalek et al. [203]. We used the direct (Figure 3-6 on the following page A) and the loop transfer method (Figure 3-6 on the next page B) (three grids per method) to screen for lipid monolayer crystals of SbpA vitrified on holey carbon films. For each grid, ~50 crystals were imaged on CCD camera, none of which diffracted beyond a resolution of 25 Å (Figure 3-6 on the following page, insets).

3.12.5 Frozen-Hydrated SbpA Crystals on Continuous Carbon Film

We next tested whether continuous carbon film would be a better option for vitrifying SbpA crystals formed on lipid monolayers. Images taken from crystals transferred to

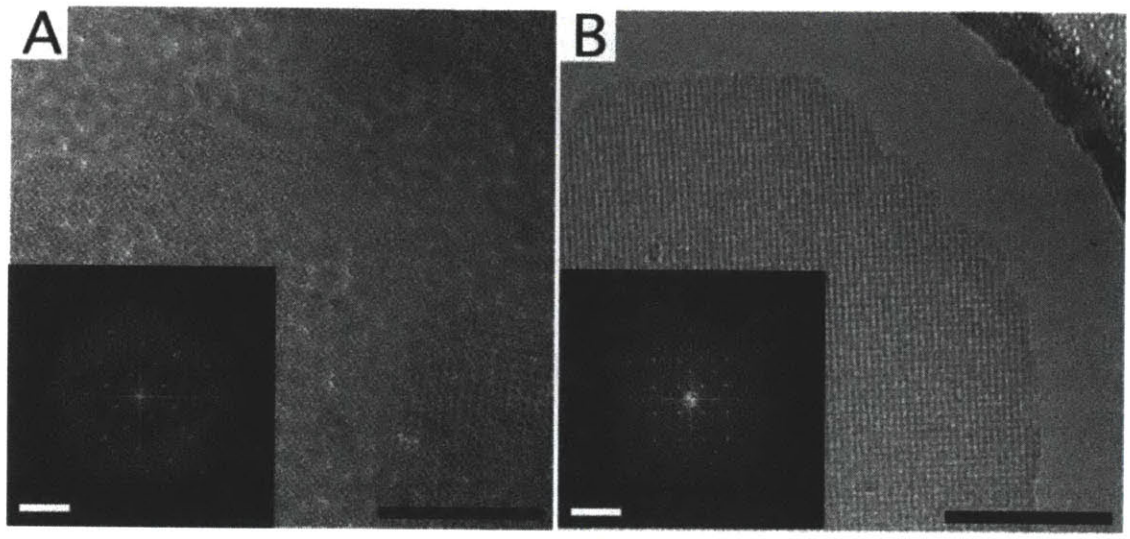


Figure 3-6: Images of Frozen-Hydrated SbpA Monolayer Crystals Prepared on Holey Carbon Grids. *(A)* SbpA crystals transferred to Quantifoil grids using the direct method. *(B)* SbpA crystals transferred to Quantifoil grids using the loop method. The Fourier transforms of these representative images shown as insets in panels *A* and *B* do not show diffraction spots beyond a resolution of 25 Å. Scale bars in the images are 100 nm; scale bars in the Fourier transform are $(6.5\ \text{nm})^{1}$.

the EM grid using the direct method also showed diffraction spots to about 25 Å, similar to images taken from crystals on holey carbon film. Using the loop rather than the direct method to transfer crystals to continuous carbon grids was an improvement as some images showed diffraction spots close to 15 Å resolution. To quantify this improvement, we prepared three grids of vitrified SbpA crystals on continuous carbon film either by direct or loop transfer. We collected 18 images from each of the six grids on film and inspected their diffraction patterns on an optical laser bench. Overall, 69% of the images taken from crystals prepared with the direct transfer did not diffract, and only 7% of the images showed diffraction spots in the resolution range from 30 to 26 Å (Figure 3-7 on the next page A). In the case of images taken from crystals prepared using the loop transfer, 39% of the images still did not show any diffraction, but 33% of the images showed diffraction spots in the resolution range from 30 to 16 Å and 7% of the images showed diffraction in the resolution range from 20 to 16 Å (Figure 3-7 on the facing page B). Using loop transfer in combination with

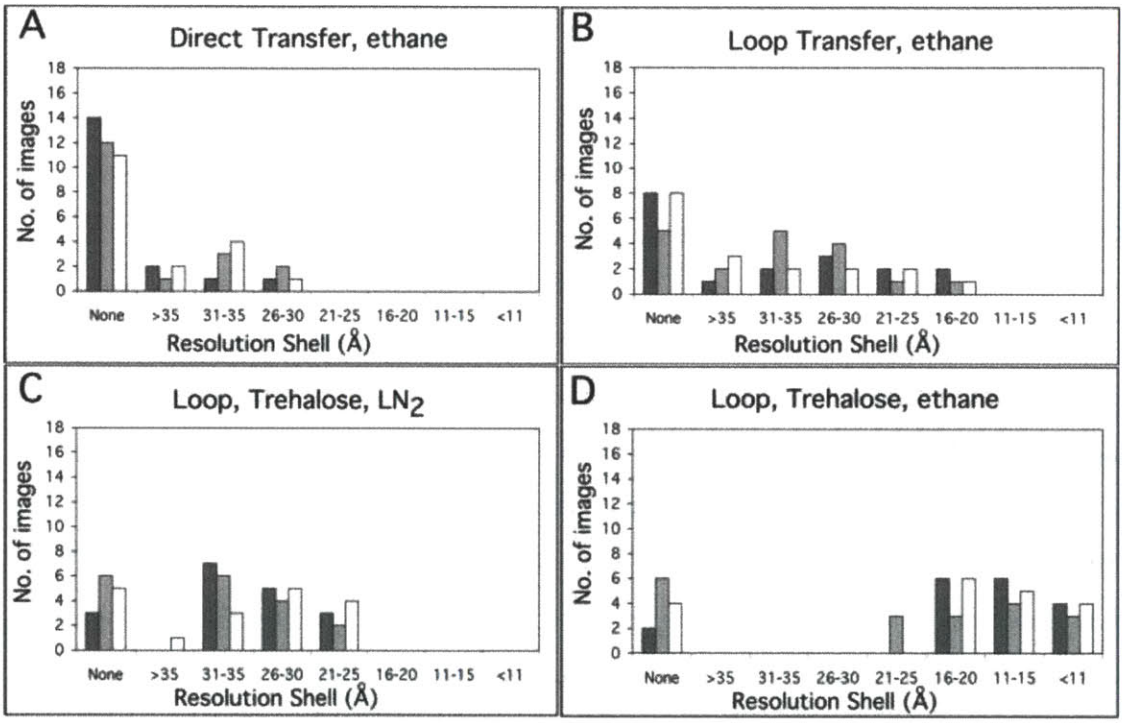


Figure 3-7: Resolution of Images Taken from Crystals Prepared on Continuous Carbon Grids by the Indicated Methods. For each preparation method, 3 grids were prepared and 18 images of SbpA crystals were taken from each grid (indicated by different grey shades). The resolution of the imaged crystals was assessed by the highest diffraction spots seen by optical diffraction. The graphs show the number of images taken from each grid that showed diffraction spots in the indicated resolution ranges.

continuous carbon film thus resulted in a modest improvement in the preservation of the SbpA crystals.

Trehalose-Embedded SbpA Crystals

Sugar-embedding, especially in trehalose [147, 158, 172], is currently considered the best preparation method for 2D crystals formed by reconstitution of membrane proteins into lipid bilayers. We therefore wanted to test whether sugar embedding could be adapted for the use with SbpA 2D crystals formed on lipid monolayers. To embed SbpA crystals in sugar we transferred the crystals to grids covered with a continuous carbon support film and allowed them to adsorb for 2 min. We then added 2 µl of 5% trehalose solution either to the front side of the grid (carbon-coated side), to which the crystals were adsorbed, or to the back side of the grid (grid-bar side). The grids were blotted and frozen in liquid nitrogen or in liquid ethane. All attempts to produce trehalose-embedded specimens using the direct transfer method failed. This is most likely due to the fact that crystals directly transferred to a grid interact with the carbon film through the hydrophobic tails of the lipid monolayer (Figure 3-3 on page 92 C). This interaction appears to be too weak to withstand the additional treatment of the grid with sugar solution, presumably causing the crystals to be sucked away during blotting or to be ripped off the grid during plunge-freezing.

We then transferred the crystals to continuous carbon grids using the loop method, in which case it is the protein array that makes contact with the carbon film (Figure 3-3 on page 92 E). When we applied sugar to the front side of the grid (carbon-coated side), many of the crystals were covered with small lipid vesicles (similar to what can be seen in Figure 3-5 on page 100 B), which caused distortions in the SbpA arrays. We found that the best way to prepare sugar embedded specimens was to use the loop method to transfer the crystals to a glow-discharged grid covered with a continuous carbon film, allow the crystals to adsorb for 2 min, add 2 µL of 5% trehalose solution onto the back side of the grid, the grid-bar side, (spreading the drop evenly so that the trehalose can diffuse into the specimen and embed the crystals), blot the grid from the back side for 10 s with Whatman #1 paper before freezing

either in liquid nitrogen or liquid ethane. The quality of specimen preservation was quantified as described above by taking 18 images from three different grids and assessing the resolution by optical diffraction. Freezing sugar embedded samples in liquid nitrogen did not produce significantly better results than vitrification in the absence of trehalose. Twenty-six percent of the images did not show any diffraction and the best images (17%) only showed diffraction spots to a resolution of about 21 Å (Figure 3-7 on page 103 C). Freezing sugar embedded samples in liquid ethane dramatically improved the preservation of the crystals. Only 22% of the images did not show any diffraction, the majority of images (72%) showed spots to a resolution better than 20 Å and 20% of the images showed diffraction spots to a resolution better than 11 Å (Figure 3-7 on page 103 D).

Projection Map of SpbA at 7 Å Resolution

Low-dose images of lipid monolayer crystals of SbpA, transferred to a continuous carbon grid with the loop technique and embedded in 5% trehalose before freezing in liquid ethane, were used to calculate a projection map. The quality of the images and the imaged crystals was assessed by optical diffraction. The best ordered crystalline areas (showing diffraction spots beyond a resolution of 10 Å) in high-quality images (with no signs of specimen drift or charging) were selected for further image processing. After unbending and correction for the CTF, CTF plots typically showed good completeness of diffraction data to ~7 Å resolution, and occasionally IQ 3 and 4 spots to a resolution of about 4 Å (Figure 3-8 on the following page A). The SbpA 2D crystals have unit cell dimensions of a = b = 133 ± 0.5 Å and $\gamma = 90°$ (n = 8) and display *p4* symmetry, determined using the program ALLSPACE [343] (Table 3.1 on page 107). The amplitude and phase data from the best five images were merged to 7 Å resolution (Table 3.2 on page 108) while applying *p4* symmetry. The final projection map of trehalose-embedded SbpA is shown in Figure 3-8 on the next page B. The projection map shows that a unit cell of the SbpA 2D crystals (outlined in black in Figure 3-8 on the following page B) contains a single tetramer, as previously described for native cell walls [212, 213].

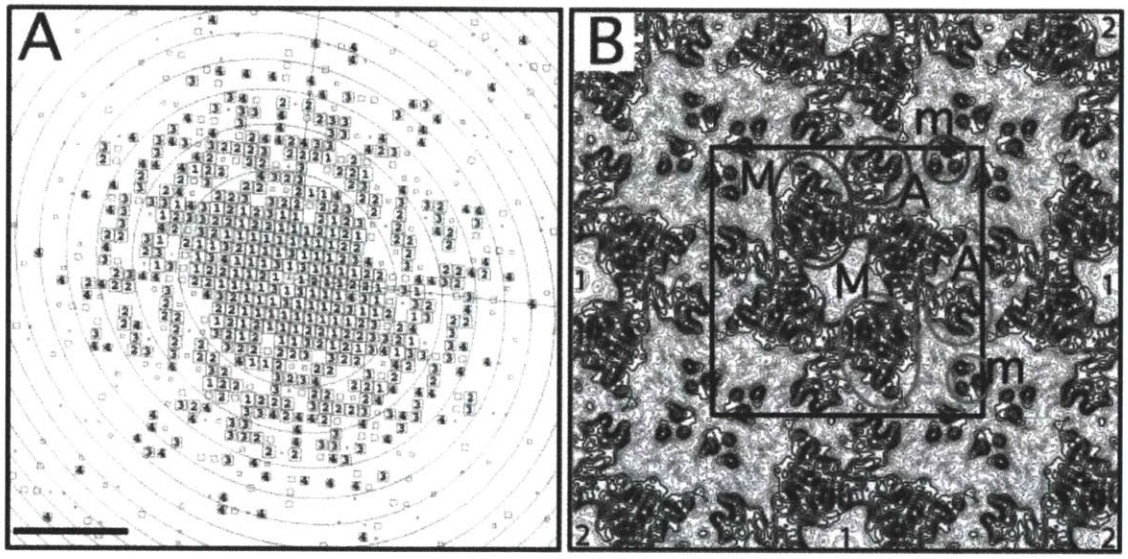


Figure 3-8: Electron Crystallographic Analysis of Trehalose-Embedded SbpA Lipid Monolayer Crystals. *(A)* CTF plot of a representative image of a trehalose-embedded 2D crystal. The circles represent CTF zero transitions and the boxed numbers depict the IQ values of the individual reflections as defined by Henderson *et al.* [154]. Scale bar is $(1.5\ \text{nm})^1$. *(B)* *p4*-symmetrized projection map of SbpA at 7 Å resolution. A unit cell with a side length of a = b = 133 Å is outlined in black. The major (M), arm (A) and minor (m) domains of two SbpA subunits are labeled. The light grey circles show the arrangement of the domains if the monomer adopted a compact organization in the 2D arrays and the dark grey circles the arrangement of the domains if the monomer adopted an extended organization. Nearest and second-nearest neighboring SbpA tetramers with respect to the central tetramer are labeled by "1" and "2", respectively.

Table 3.1: Internal Phase Residuals of All Possible Two-Sided Plane Groups Using a Representative Image. Internal phase residuals were determined using the program ALLSPACE [343] from spots of IQ1 to IQ5 to 7 Å resolution. Only plane groups compatible with the SbpA lattice are shown.
[a] Phase residual versus other spots (90° random).
[b] Target residual based on statistics taking Friedel weight into account.
[c] Note that in space group $p1$ no phase comparison is possible, so the numbers given here are theoretical phase residuals based on the signal-to-noise ratio of the observed diffraction spots.
[d] Within 20% of target residual.

Two-Sided Plane Group	Phase Residual (Degrees)[a]	Number of Comparisons	Target Residual (Degrees) [b]
$p1$	27.8[c]	434	
$p2$	40.8[d]	217	40.8
$p12_b$	73.7	113	29.3
$p12_a$	74.0	111	29.0
$p12_1b$	76.6	113	29.3
$p12_1a$	72.4	111	29.0
$c12_b$	73.7	113	29.3
$c12_a$	74.0	111	29.0
$p222$	59.0	441	34.2
$p222_1b$	58.4	441	34.2
$p222_1a$	62.2	441	34.2
$p22_12_1$	69.3	441	34.2
$c222$	59.0	441	34.2
$p4$	30.0[d]	525	33.2
$p422$	51.3	998	30.6
$p42_12$	61.6	998	30.6

Table 3.2: Phase Residuals in Resolution Shells.
$\Delta\alpha_c$, Difference between the symmetry-imposed phase of 0° and 180° and the observed combined phase. 45° is random.

Resolution Shell (Å)		Number of	Mean Value of	Standard Error	Mean Figure
From	To	Phases	$\Delta\alpha_c$ (Deg.)	(Deg.)	of Merit
200.0	20.0	34	12.7	2.1	0.95
20.0	12.0	62	22.7	2.9	0.86
12.0	9.0	74	25.0	2.5	0.82
9.0	7.0	112	36.5	2.3	0.65
7.0	6.0	105	43.6	2.5	0.52

The SbpA tetramer has a complex projection structure with each subunit consisting of three domains as described previously [212, 213]. The four major domains (labeled M in Figure 3-8 on page 106 B) form a square and constitute the core of the SbpA tetramer. The four major domains surround a square low-density area in the center of the tetramer. Four smaller arm domains (labeled A in Figure 3-8 on page 106 B) extend from the sides of the core tetramer. The arm domains make contact with the major and arm domains of one subunit from each of the four neighboring tetramers. Each SbpA tetramer thus forms strong connections to the four nearest neighboring SbpA tetramers (labeled "1" in Figure 3-8 on page 106 B) through interactions involving the major and arm domains. The minor domain of the SbpA subunit (labeled m in Figure 3-8 on page 106 B) does not show an obvious connection to the major and arm domains. The light and dark grey circles thus indicate the two possibilities how the three domains may be arranged in an SbpA subunit. Minor domains originating from four different SbpA tetramers also form a tetramer along the four-fold axis at the corners of each unit cell (Figure 3-8 on page 106 B). Each minor domain thus mediates connections with two of the nearest neighboring SbpA tetramers and a second-nearest neighboring SbpA tetramer (labeled "2" in Figure 3-8 on page 106 B).

3.13 Discussion

Advances in specimen preservation have played a pivotal role in obtaining high-resolution structures for membrane proteins by electron crystallography of 2D crystals. Sugar embedding was originally introduced when Unwin and Henderson preserved purple membranes in glucose for electron microscopic imaging [340], resulting in the first 3D density map of a membrane protein [153] and later in the first atomic model of a protein based on electron crystallographic data [155]. The sugar serves both as non-volatile replacement for water and as cryo-protectant in the freezing of the crystals. In addition, because sugar embedding involves much more extensive drying of the sample as compared to sample vitrification, specimens prepared in sugar tend to be flatter and better suited for collecting data from tilted specimens. Tannic acid was subsequently used to preserve 2D crystals of plant light-harvesting complex II [98, 346] and trehalose has been used to preserve a number of 2D crystals [158, 172, 252]. Trehalose is the sugar expressed by organisms in response to cold shock [182] and it has characteristics that make it particularly well suited for the preservation of protein structure upon freezing [77, 158]. The main purpose of this study was thus to evaluate whether trehalose embedding could be adapted for specimen preparation of 2D crystals grown on lipid monolayers and whether it would preserve the structure better than the currently used techniques. Sugar embedding has been used before to prepare specimens of streptavidin 2D crystals grown on lipid monolayers. Kubalek et al. [203] used the direct transfer method to apply streptavidin crystals to reticulated carbon grids and used 1% glucose in combination with freezing in liquid nitrogen to preserve the crystals. Electron diffraction patterns of untilted crystals prepared in this way showed diffraction spots to a resolution of 2.8 Å. Avila-Sakar and Chiu [37] also prepared streptavidin lipid monolayer crystals by direct transfer to reticulated carbon grids, but they froze the grids in liquid ethane without prior addition of sugar. Using this preparation method the authors could collect electron diffraction patterns and images that allowed them to calculate a projection map at 3 Å resolution. The authors also reported that the vitrified crystals

were not sufficiently flat to collect data from tilted specimens, but that they could record high-resolution electron diffraction patterns from glucose-embedded crystals at a tilt angle of 50° [37]. These studies thus not only report the highest resolution data obtained with lipid monolayer crystals to date, but also suggest that sugar embedding may indeed also be the best specimen preparation procedure for lipid monolayer crystals to collect data from tilted specimens.

Our results differ from those obtained by Kubalek et al. [203], as we were unable to produce sugar-embedded specimens using the direct transfer method for our SbpA lipid monolayer crystals. The difference may be explained by the fact that we used continuous carbon film rather than reticulated carbon film and the associated differences in the way the sugar was applied to the specimen. Instead, we used the loop transfer technique developed by Asturias and Kornberg [36] to adsorb our monolayer crystals to continuous carbon grids. Since with this technique the protein array interacts with the carbon film rather than the hydrophobic lipid tails of the monolayer, the crystals are more firmly attached and can thus withstand the sugar embedding procedure. A surprising finding of our study is that trehalose embedding in combination with freezing in liquid nitrogen (Figure 3-7 on page 103 C) did not improve the preservation of the SbpA crystals compared to vitrified samples (Figure 3-7 on page 103 B), while trehalose-embedded specimens plunge-frozen in liquid ethane showed a significantly better preservation of the crystals (Figure 3-7 on page 103 D). The reason for this difference is not clear, since the trehalose should function as cryo- protectant and make the freezing rate irrelevant for specimen preservation. Nevertheless, the difference between freezing the monolayer crystals in liquid nitrogen versus freezing them in liquid ethane was substantial and reproducible. Our specimen preparation protocol, which consists of loop transfer of the monolayer crystals to continuous carbon grids, embedding with 5% trehalose and subsequent freezing in liquid ethane, allowed us to collect images and calculate a projection map of SbpA at 7 Å resolution (Figure 3-8 on page 106). This resolution is low compared to a resolution of 3 Å obtained with streptavidin monolayer crystals [37, 203]. Streptavidin is, however, a much smaller protein (MW 15 kDa) and forms highly ordered 2D crystals with

a smaller unit cell (a = b = 82.3 Å). Streptavidin 2D crystals are therefore probably more rigid than 2D arrays formed by the larger SbpA tetramer and hence less susceptible to damage introduced by the specimen preparation procedure. This notion is supported by the fact that data of comparable resolution could be collected from vitrified [37] and glucose-embedded [203] streptavidin crystals. This is clearly not the case for SbpA crystals, since images of specimens prepared by direct transfer to holey carbon grids and subsequent freezing, in contrast to our images collected from trehalose-embedded crystals, never diffracted beyond a resolution of 25 Å. Although we cannot be certain, it is likely that the resolution of our SbpA projection map is not limited to a resolution of 7 Å by the specimen preparation method but by the order of the SbpA arrays themselves. We predict that images of streptavidin crystals prepared by the trehalose-embedding protocol presented in this manuscript would also diffract to at least 3 Å resolution, although this prediction needs experimental testing.

Our 7 Å projection map of SbpA is a substantial improvement compared to previous structural studies on SbpA arrays, which were all limited to a resolution of about 20 Å [24, 212, 213, 286]. Our higher-resolution projection map corroborates the previously published low-resolution maps [24, 212, 213]. The first negatively stained projection map of SbpA already revealed the three domains of an SbpA subunit [24], which were subsequently named major domain, arm domain and minor domain [212]. These three domains became more clearly defined in a later 3D reconstruction from negatively stained SbpA crystals [213]. The complex 3D structure of the SbpA tetramer makes it difficult to interpret our projection map in terms of secondary structure elements despite the resolution of 7 Å, which should resolve α-helices. Only the minor domain shows distinct structural features and is resolved in our map into three circular densities (Figure 3-8 on page 106 B). The diameter of these three densities of about 8 Å would be consistent with three helices forming a bundle and running perpendicular to the plane of the lipid monolayer.

Neither the previous low-resolution maps nor our 7 Å projection map reveal how the three domains are connected in an SbpA subunit. One possibility is that the

three domains form a compact arrangement, where the major, arm and minor domains form a triangle (light grey circles in Figure 3-8 on page 106 B). Alternatively, the three domains could form an extended, linear arrangement as indicated by the dark grey circles in Figure 3-8 on page 106 B. Although the precise domain organization of SbpA is not clear, similar architectures are found in other tetragonal surface layers, as for example in the S-layer from *Sporosarcina ureae* (*S. ureae*) [120]. Higher resolution information on the 3D structure will be needed, however, to elucidate the exact building plan of these tetragonal S-layers. Knowledge of the structure will be crucial in attempts to engineer recombinant S-layer proteins that form arrays with designed properties. S-layers have potential for use in the fabrication of biotemplated energy generating devices such as batteries [228, 258] and solar cells [101], where precise engineering of the S-layer structure can create a scaffold for the growth and arrangement of nanomaterials. A 3D reconstruction at a resolution of 10 Å or better will require flat specimens that can be used to collect EM data from tilted specimens. Since sugar embedding produced flatter specimens of streptavidin crystals grown on lipid monolayers [37], the trehalose embedding technique presented here may lay the foundation to determine 3D structures of lipid monolayer crystals of SbpA as well as other proteins at resolutions better than 10 Å.

3.14 Contributions

Norville, Belcher, and Knight wished to explore the structure of the SbpA after assembly into 2D crystals at the highest resolution possible. Norville, Kelly, Belcher, Knight, and Walz designed the experiments. Norville, Kelly, and Walz wrote the manuscript.

3.15 Acknowledgements

This work was supported by the Army Research Office Institute of Collaborative Biotechnologies (to AMB), the Army Research Office Institute of Soldier Nanotech-

nologies (to AMB), the David and Lucile Packard Foundation (to AMB), DARPA/ONR N00014-01-1-1060 (to TK), Semiconductor Research Corporation task ID 1267.0001 (to TK), a National Science Foundation Graduate Research Fellowship (to JN), a Vinton Hayes Graduate Research Fellowship (to JN), and a MIT/Merck Graduate Research Fellowship (to JN). The molecular EM facility at Harvard Medical School was established with a generous donation from the Giovanni Armenise Harvard Center for Structural Biology and is supported by National Institutes of Health Grant GM62580 (to TW). Any opinions, findings, conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Chapter 4

Cells-on-Paper for Modular Design of Cellular Ensembles

In Chapter 1 on page 19 I described how the interfaces of shipping containers remain the same, even if their external contents are altered, and how the shipping containers can be brought together using those interfaces. I also discussed how homeobox genes partition the cells that make up a multicellular biological life forms into spatial compartments, which then develop according to downstream processes. The Hox genes divide arthropods and chordates into spatial containers along the anteriorposterior axis [78].

The cells-on-paper system provides a way to compose spatial compartments of cells simply by stacking different layers of cell-embedded paper and interleaving the paper layers with cell-impermeable membranes. One can mix and match the contents of different compartments to create new systems. One can change the positions of compartments relative to one another. One can change the types of signals that can traverse the cell-impermeable, yet chemically permeable interfaces.

4.1 Structure and Contents

In this chapter I describe a method to seed bacterial cells in paper layers, stack the layers, and isolate the different strains using isoporous membranes. Communication

signals that diffuse between the structures can be designed to elicit a behavior in a different spatial regions. Currently the top layer of a paper stack will be oxygenated, while lower layers are anoxic. The paper platform can be used to construct synthetic systems using multiple cell strains or to study the properties of the cells that make up a biofilm.

This chapter contains a single manuscript. Chapter 4 uses material from the document: Julie E. Norville, Steven Pennybaker, Ratmir Derda, Matthew R. Lockett, Cory Li, George Whitesides, Angela M. Belcher, and Thomas F. Knight, Jr. "Paper supported polycultures of synthetically programmed bacteria." An abstract describing the manuscript is included in the final dissertation summary in Section 5.1.4 on page 140. A list of abbreviations is included in Section 1.5 on page 48.

4.2 Introduction

4.2.1 Studying Microbial Communities

Most microbes exist as 3D communities (e.g., biofilms) in their natural habitat. The evolution and stability of these multicellular structures is of interest as they have been shown to play an important role in human health [150]. In addition, mature biofilms contain metabolic environments, physiological responses (e.g., increased resistance to antimicrobials), and genetic adaptations that are absent in populations of planktonic cells grown in liquid media under laboratory conditions [328]. In addition, the small molecule or peptide-based intraspecies and interspecies communication signals between bacteria can influence the formation, behavior, and maintenance of bacterial biofilm communities [46, 136, 177, 177, 315]. Long-range communication between individual gram-negative bacteria within a community occurs via the diffusion of signaling molecules (e.g., HSL and peptides) [148].

There exists a need for simpler strategies [275, 326] to study the formation and maintenance of biofilms that resemble naturally occurring polycultures as well as to construct structured ensembles of synthetic bacteria. An ideal platform for biofilm

studies would allow one to: (i) construct a defined 3D community; (ii) easily analyze the spatial and temporal behavior and composition of the community; (iii) isolate bacteria from defined locations within the community; (iv) control the gradients of signaling molecules within the community. We thus present here a simple approach to construct and study bacterial ensembles by exploring the effect of small molecule communication within a bacterial community using 3D paper scaffold.

4.2.2 3D Culture of Bacterial Cells in Paper Structures

Here we propose to use a paper-platform to study the effects of intercellular signaling in natural and synthetic systems. We have described, previously, a means to mimic tissues by positioning mammalian cells within a 3D construct composed of a stack of ordinary filter paper using supporting slabs of gel [105]. We propose that co-stacking layers of bacterial cells, analogous to the mammalian cell culture system, will result in a simple and versatile tool for generating and analyzing 3D communities of bacteria. Destacking the layers of paper isolates slices containing viable cells from the culture and allows one to the examine the behavior of the cells inside these 3D cultures as a series of 180 μm-thick slabs. The ability to destack each layer will afford a simple alternative to visualizing cells within the culture; previous methods have relied on physical sectioning techniques (e.g., cryosectioning and laser capture microsectioning) [328], which are difficult and time consuming and do not provide a way to study living cells. Although the paper thickness used here (180 μm) is much larger than the 10-50 μm thick biofilms typically formed by *E. coli* [174, 277], the different sheets partition the film into an oxygenic and anoxic layer (as 175 μm is the depth of oxygen diffusion [328] providing a simple way to examine cellular composition and behavior in each environment.)

E. coli was chosen for this system because of its use as a model organism throughout microbiology [111] and in studying cell growth [232], signaling patterns [47], and biofilms [231]. By destacking the layers of paper, we have examined the behavior of cells inside the 3D-platform. We tested the paper-platforms ability to allow for the creation and perturbation of 3D chemical gradients and intercellular signal response.

We tested this ability by characterizing a synthetic intercellular signaling model. The communication between two *E. coli* strains: HSL producing "Sender" bacteria and HSL-responsive "Receiver" bacteria, were used as our signaling model. The characterization of this paper-based platform illustrates how it can be used to test natural or synthetic systems that use chemical gradients, interspecies signaling, and cellular signaling. While researching the platforms versatility as a way to study cellular signaling, we conducted tests that suggest the platform develops an oxygen gradient, an important physiological characteristic of natural cellular aggregates [328]. We propose that this paper-based platform can be used as a valuable technique to study 3D bacterial co-culture response to intercellular signaling and chemical perturbations as well as construct synthetic systems using multiple strains.

4.3 Materials and Methods

4.3.1 Paper Substrates and Substrate Preparation

20 cm x 20 cm sheets of grade # 1 chromatography paper (Whatman), with a porosity of 11 μm and a thickness 180 μm, were cut to standard sizes using the following method. In preparation for trimming, sheets were mounted with tape on a paper and an acrylic-padding layer. The mounted chromatography paper was then placed in a Universal Laser Systems X2-660 laser cutter programmed to excise 6.5 cm x 4 cm rectangles, 2 cm x 2 cm squares, and isosceles right triangles with 8 mm legs from the sheet, using settings designed to engrave cork to a depth of 250 μm. For sterilization and ash removal, the excised shapes were submerged in a glass petri dish in a beaker containing 1 L of purified and deionized (Milli-Q) water and autoclaved. Water was then drained from the petri dish and the excised sheets were dried in a laminar flow hood under a UV light for 24 h.

4.3.2 "Sender" Vectors

The "Sender" strain contained two plasmids that each constitutively expressed a protein. The ampicillin resistant plasmid for the "Sender" was the vector pSND-1 [356]. Vector pSND-1 contains a LuxI gene that encodes for a 3-oxohexanoyl-homoserine (3OC6HSL) synthetase. The HSL producing machinery was borrowed from the quorum-sensing organism *Vibrio fischeri* (*V. fischeri*) [356]. The chloramphenicol resistant plasmid, pSB3C5, contained the BioBrick part BBa_I7001 (http://partsregistry.org/wiki/index.php?title=Part:BBa_I7001), encoding enhanced yellow fluorescent protein (EYFP). Both vectors were cut with the restriction enzymes EcoRI and SpeI, and the protocol from the BioBrick™ Assembly Kit [20] was used to transfer the part BBa_I7001 to the pSB3C5 vector.

4.3.3 "Receiver" Vectors

The "Receiver" bacteria contained two plasmids: the ampicillin resistant vector pRCV-3[356], that had been engineered to express GFP in the presence of HSL [356] (i.e., the HSL-sensing machinery also borrowed from *V. fischeri*) and the chloramphenicol resistant vector pSB4C5 containing BioBrick part BBa_J04450 (http://partsregistry.org/Part:BBa_J04450), which encodes RFP.

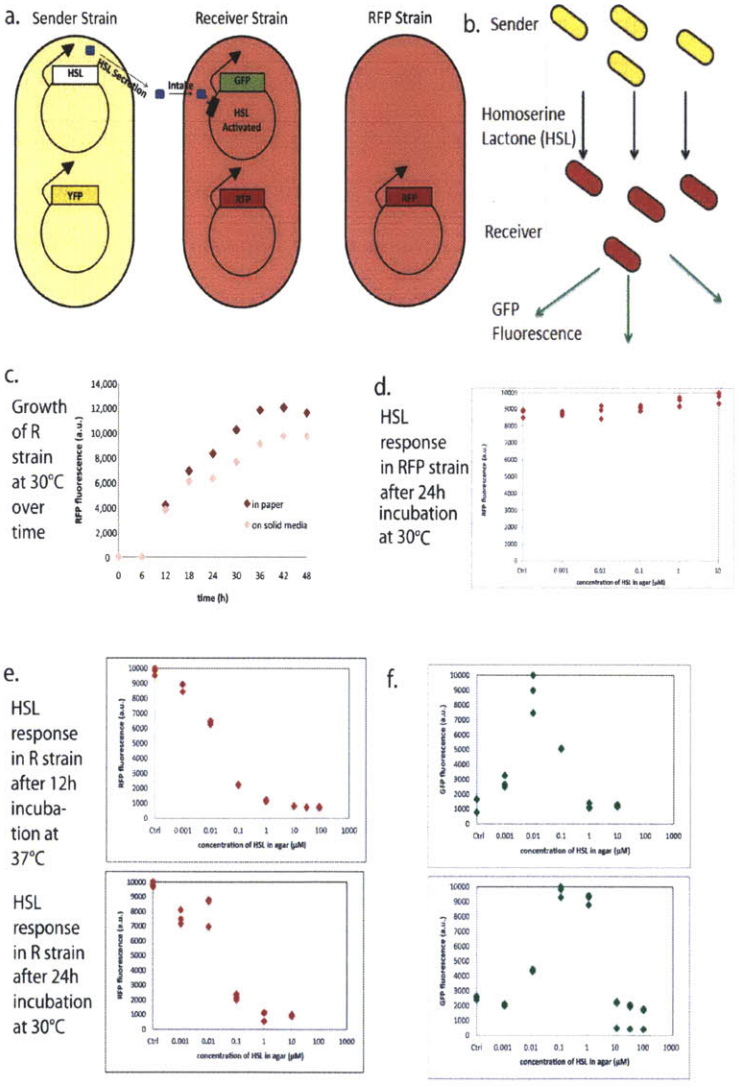


Figure 4-1: "Sender," "Receiver," and "RFP" bacteria and "Receiver" Fluorescence *(a)* The genetically engineered *E. coli* strains used. The "Sender" (S strain) contains two plasmids: one producing HSL and one producing YFP. The "Receiver" (R strain) contains the RFP-producing plasmid and the HSL-activated GFP-producing plasmid *(b)* Representation of "Sender" (S) and "Receiver" (R) strain interaction. *(c)* RFP fluorescence of "Receiver" (R) strain as a function of time, both in paper and on solid media. *(d)* RFP expression by non-HSL responsive "RFP"-strain as a function of HSL-in-agar concentration, measured via fluorescence intensities on a Typhoon 9400 Variable Mode Imager. *(e)* Expression of RFP by "Receiver" bacteria as a function of HSL-in-agar concentration after 12 h incubation at 37°C and 24 h incubation at 30°C. Fluorescence intensities measured on a Typhoon 9400 Variable Mode Imager.

4.3.4 "RFP" Strain Vectors

The "RFP" strain contained only one plasmid: the chloramphenicol resistant vector pSB4C5 containing BioBrick part BB_J04450. This was the same vector that was used to express RFP in the "Receiver" bacteria.

4.3.5 Bacterial Strains and Culture

10-β chemically competent *E. coli* (New England Biolabs) were transformed using the instructions from the manufacturer. Creation of either the "Sender" or "Receiver" bacteria required fresh cotransformants that contain two plasmids (see Figure 4-1 on the preceding page), one encoding ampicillin resistance and the other encoding chloramphenicol resistance. Thus, for all experiments the "Sender" and "Receiver" bacterial strains were maintained on LB agar plates or cultured in LB medium that contained 100 μg/ml ampicillin and 35 μg/ml chloramphenicol. Vectors were obtained or modified from plasmids obtained from the Knight laboratory or the Registry of Standard Biological Parts (www.partsregistry.org). 1:100 dilutions of the "Sender" and "Receiver" bacterial strains were individually grown in LB medium at 37°C until a cell density of 0.6 at OD_{600} was reached before being used in experiments.

4.3.6 Bacterial Cultures in Paper

"Sender" and "Receiver" cultures that had each reached 0.6 OD_{600} were diluted 1:100 in LB media containing the appropriate antibiotics. Paper was seeded with bacteria, manipulated using sterile forceps, and assembled into stacks on a sterile agar plate containing 25 mL of LB agar. "Sender" seeded triangles and "Receiver" seeded rectangles were separated using a 0.2 μm isopore membrane (SPI) for most experiments. A sterile razor blade was used to separate the stacked sheets before they could be manually removed.

All cultures were grown at either 37°C for 12 h or 30°C for 24 h. Water was placed in the incubator to maintain humidity. Bacteria-seeded paper was scanned directly after removal from incubator, unless stored for visual observation. Samples

for visual observation were placed in 4°C refrigerator. The images were obtained by scanning the paper either on agar plates or directly on glass using a fluorescent gel scanner (Typhoon 9400 Variable Mode Imager). The excitation laser used for YFP and RFP had a wavelength of 532 nm whereas the laser for GFP measurements had a wavelength of 488 nm. The emission wavelengths measured and PMT settings were 555 nm and 300 V for YFP, 520 nm and 500 V for GFP, and 610 nm and 300 V for RFP. The resolution was 50m unless otherwise specified. Data were analyzed using Matlab and ImageJ.

4.3.7 Fluorescence Intensity Calibration Curves

The "Receiver" was cultured to an OD_{600} of 0.6 and centrifuged at 3500 g for 15 minutes. The x10, x100, and x1000 cellular concentrations were obtained by re-suspending the pellet in the appropriate volume of LB media. The x1, OD_{600} 0.6 culture was used directly. For each of the three replicates, 100 μL of each concentration was absorbed into a paper square. The samples were placed on agar plates for transport to the gel imager and then removed and scanned on glass.

4.3.8 Growth Curves

Two separate plates for each timepoint were created. One plate contained four squares inoculated with the 1:100 dilution of OD_{600} 0.6 "Receiver." The other plate had nine 5μL drops of 1:100 dilution of OD_{600} 0.6 "Receiver" applied directly to it using a 20 μL pipetman. Beginning at the 0 h timepoint, a plate of each type was harvested and scanned every 3 h until 24 h timepoint was reached.

4.3.9 Bacterial Response to HSL

Bacterial response to HSL concentration was determined by preparing LB plates containing the antibiotics and various amounts of HSL. The HSL used was 3OC6HSL, which is also known by the name N-(β-Ketocaproyl)-L-homoserine lactone (Product K3007, Sigma-Aldrich). 3OC6HSL was added to warm agar solution and mixed before

plating and solidification. Two plates at concentrations of 0.001 μM, 0.01μM, 0.1 μM, 1 μM, 10 μM, and 100 μM HSL were obtained.

4.3.10 Cell Counts in Paper

After destacking a cultured stack of "Receiver" containing paper, each rectangle was placed in a falcon tube containing 10 mL sterilized LB and vortexed. Dilutions were then made of the LB solution and plates onto ampicillin and chloramphenicol containing solid media. The plates were allowed to grow overnight and fluorescence scanning used to detect colonies.

4.4 Results

4.4.1 Constructing and Studying 3D Bacterial Communities in Paper Structures

3D polycultures of bacteria were formed by stacking individual layers of paper containing a clonal culture of *E. coli* on top of one another. The 3D structures were then placed atop solid LB media for further culturing. Capillary forces held the individual paper layers together. We selected *E. coli* as a model bacterial system to demonstrate the feasibility of culturing a functional 3D bacterial community.

4.4.2 Bacterial Cultures in Paper: The Strains Used to Create a Synthetic Signaling System

To characterize the usefulness of the paper platform to study composed natural or synthetic systems as well as intracellular signaling, we used a system of three bacterial strains. Strain construction was described in the Materials and Methods section. One strain, the "Sender" bacteria contained two plasmids: a plasmid that produced (or "sent") HSL and a plasmid that produced yellow fluorescent protein (YFP). A second strain, the "Receiver" also contained two plasmids: a plasmid that was engineered to

produce GFP [47] as a response to HSL (or "received" the signal) and a plasmid that produced red fluorescent protein (RFP). A third strain, the "RFP" strain contained only the RFP-producing plasmid that was previously used in the "Receiver" strain. A representation of the plasmid system, three strains used, and the HSL interaction between the sender and the receiver plasmid system can be found in Figure 4-1 on page 120 a, and a diagram summarizing the "Sender" (S) and "Receiver" (R) strain interaction expected can be found in Figure 4-1 on page 120 b.

Growth Curves

In order to determine if the growth curves of paper-seeded bacteria were similar to the growth curves of bacteria grown in more conventional ways: expression of RFP by the "Receiver" strain on both paper and agar was obtained (Figure 4-1 on page 120 c). In this experiment, paper squares inoculated with or spots containing a 1:100 concentration of 0.6 OD_{600} "Receiver" strain (10^6 cfu/ml) was placed on agar plates and scanned for RFP fluorescence every 6 h over the course of 48 h at 30°C. RFP expression appeared to level off in both on-paper and on-agar samples after approximately 36 h, suggesting that the bacteria in both conditions reached stationary growth at this time.

Response of the "RFP" Strain to HSL

To provide a control for HSL toxicity and to determine whether HSL concentration directly affects the viability or fluorescence of the "Receiver" bacteria, we tested the response of the "RFP" strain to HSL. The "RFP" strain has the same RFP-producing plasmid as the "Receiver" bacteria, but lacks an HSL responsive plasmid. In this experiment, (Figure 4-1 on page 120 d), the "RFP" strain was incubated for 24 h at 30°C under the same conditions described previously. The results indicate that the presence of HSL did not induce toxicity or interfere with RFP production in the non-HSL responsive "RFP" strain.

Response of the "Receiver" to HSL

We originally intended to use RFP fluorescence to test for the relative amount of "Receiver" bacteria at a certain place within the platform and to use GFP fluorescence to examine the response to HSL at that position within the platform. In order to determine the "Receiver's" response as a function of HSL concentration, we exposed sheets of paper containing a 1:100 concentration of 0.6 OD_{600} "Receiver" bacteria (~10^6 cfu/ml) to agar plates containing increasing concentrations of HSL. Cultures were incubated at either 30°C for 24 h or 37°C for 12 h to determine whether temperature affects bacterial response to HSL. The GFP and RFP fluorescence were measured after the designated period (Figures 4-1 on page 120 e and f).

The monitoring of RFP production in the "Receiver" relative to HSL had interesting results. A slight decrease in RFP fluorescence was seen at lower levels of HSL, while a sharp decrease in RFP fluorescence was observed at both incubation temperatures at 0.1 μM (Figure 4-1 on page 120 d). The decrease in RFP fluorescence happened at lower HSL concentrations than the decrease in GFP fluorescence. Taken together with the results of the "RFP" strain control (Figure 4-1 on page 120 d), these results suggest that a behavior different than just HSL toxicity was causing a decrease in fluorescence at high HSL concentrations. While we hypothesize that the decrease in fluorescence is caused by cells dying or being unable to grow due to all of their resources being used for protein expression, more experiments will be needed to prove the accuracy of this hypothesis.

We next sought to investigate GFP expression in the "Receiver" relative to HSL concentration. Monitoring GFP expression, we determined that after incubation at 30°C, the response of the "Receiver" to HSL increased with increasing HSL concentrations up to 0.1 μM. After incubation at 37°C, the response increased up to 1 μM HSL (Figure 4-1 on page 120 f). A decline in GFP production at higher HSL concentration was observed at both temperatures. These results suggest a slightly varied "Receiver" response to HSL concentrations at different temperatures.

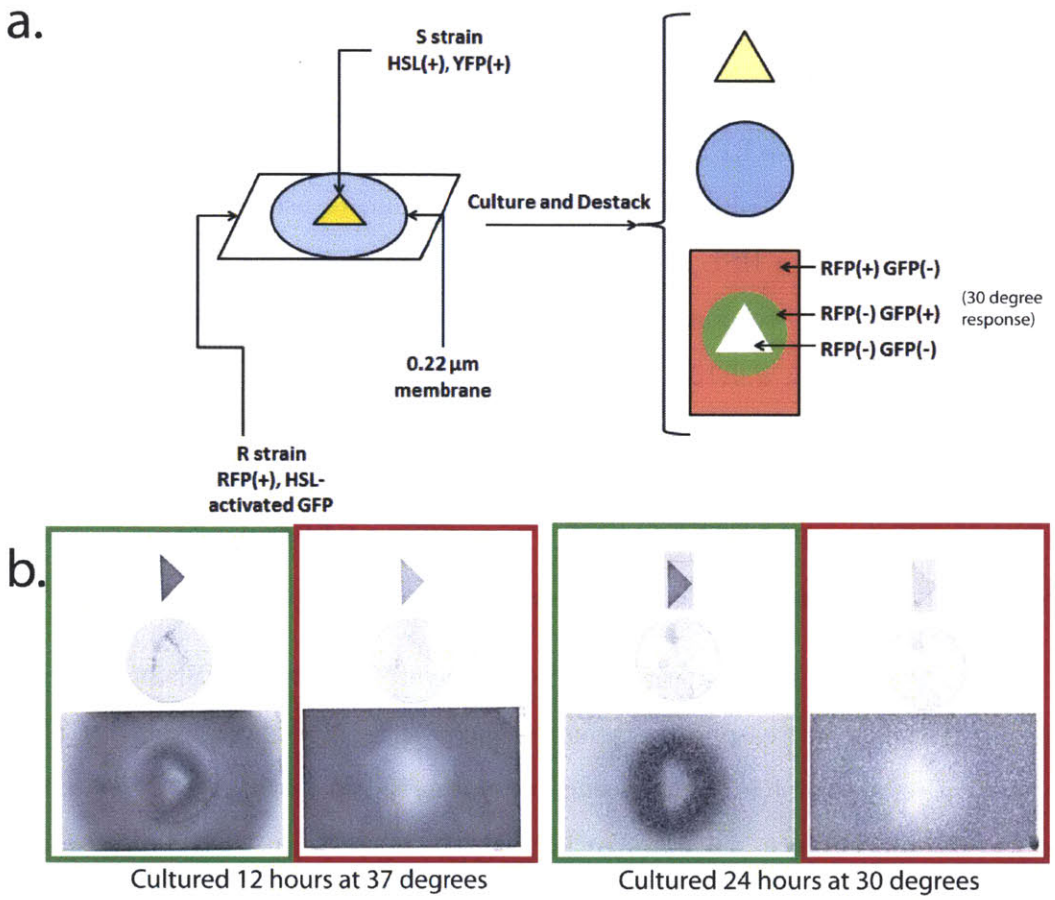


Figure 4-2: Binary Co-Culture Experiments *(a)* Scheme of bi-layer setup and pictorial representation of results. *(b)* Results from culturing bi-layer cultures at 37 °C for 12 h and at 30°C for 24 h. GFP and RFP fluorescence are highlighted by green and red rectangles, respectively.

4.4.3 *E. coli* Communication in Binary Co-Cultures – Characterization of the Response to a 2D Signaling Gradient

Once the response to HSL by the "Receiver" in the conditions above was established, we wanted to determine whether intracellular HSL-signaling between the "Sender," which produces HSL, and the "Receiver," which responds to HSL, yields a patterned behavior of signal response. We wanted to first isolate and then observe a 2D HSL-response gradient; therefore, we created "binary co-cultures," which are so-named due to their two-layered nature. The co-cultures were created by stacking two pieces

of 180 μm thick paper; one containing "Sender" bacteria and one containing "Receiver" bacteria. The "Receiver" was seeded onto 6.5 cm X 4 cm rectangles and the "Sender" was seeded onto small triangles (a 11.3 mm hypotenuse right triangle) at a concentration of 1:100 0.6 OD_{600}. The triangular shape of the "Sender"-containing paper simplified visualization of the fluorescence response. The cultures were stacked with the "Sender"-containing paper either above or below the "Receiver"-containing rectangle, with a 0.22 μm membrane placed in between the layers to prevent transfer of the strains between the paper sheets. The stacks were then cultured on a petri dish. We expected a region of GFP fluorescence radiating outwards from the "Sender" triangle, which would indicate the "Receiver" strain's response to HSL. In order to determine the effect of temperature on HSL gradients, the cultures were incubated at either 30 °C or 37°C for 24 h and 12 h, respectively. The platforms were then destacked and scanned for RFP and GFP fluorescence (Figure 4-2 on the facing page a).

We hypothesized that the region of highest HSL concentration in the rectangle would be directly under the triangle, due to this area's proximity to the HSL producing bacteria. The RFP fluorescence was lowest in the high HSL region, with a similar response in cultures at both temperatures (Figure 4-2 on the preceding page b). The response is not surprising, given the decrease in RFP fluorescence in the "Receiver" at high HSL concentrations (Figure 4-1 on page 120 e). A distribution of GFP expression was observed at both 30°C and 37°C cultured "Receiver" strains. At both temperatures, a circular pattern of high fluorescence was observed (Figure 4-2 on the preceding page b). The temperature seemed to affect the HSL response patterns, however. Under the 37°C incubation temperature, the radius of GFP expression was larger than the corresponding radius on the 30°C cultured sample. Also observed was a triangular area of no fluorescence, corresponding to where the "Sender" triangle had been placed. This triangular pattern was observed in both "Sender"-on-top and "Sender"-on-bottom configurations, indicating that the primary cause is not potential oxygen deprivation due to the presence of the "Sender" layer on top of the "Receiver" layer.

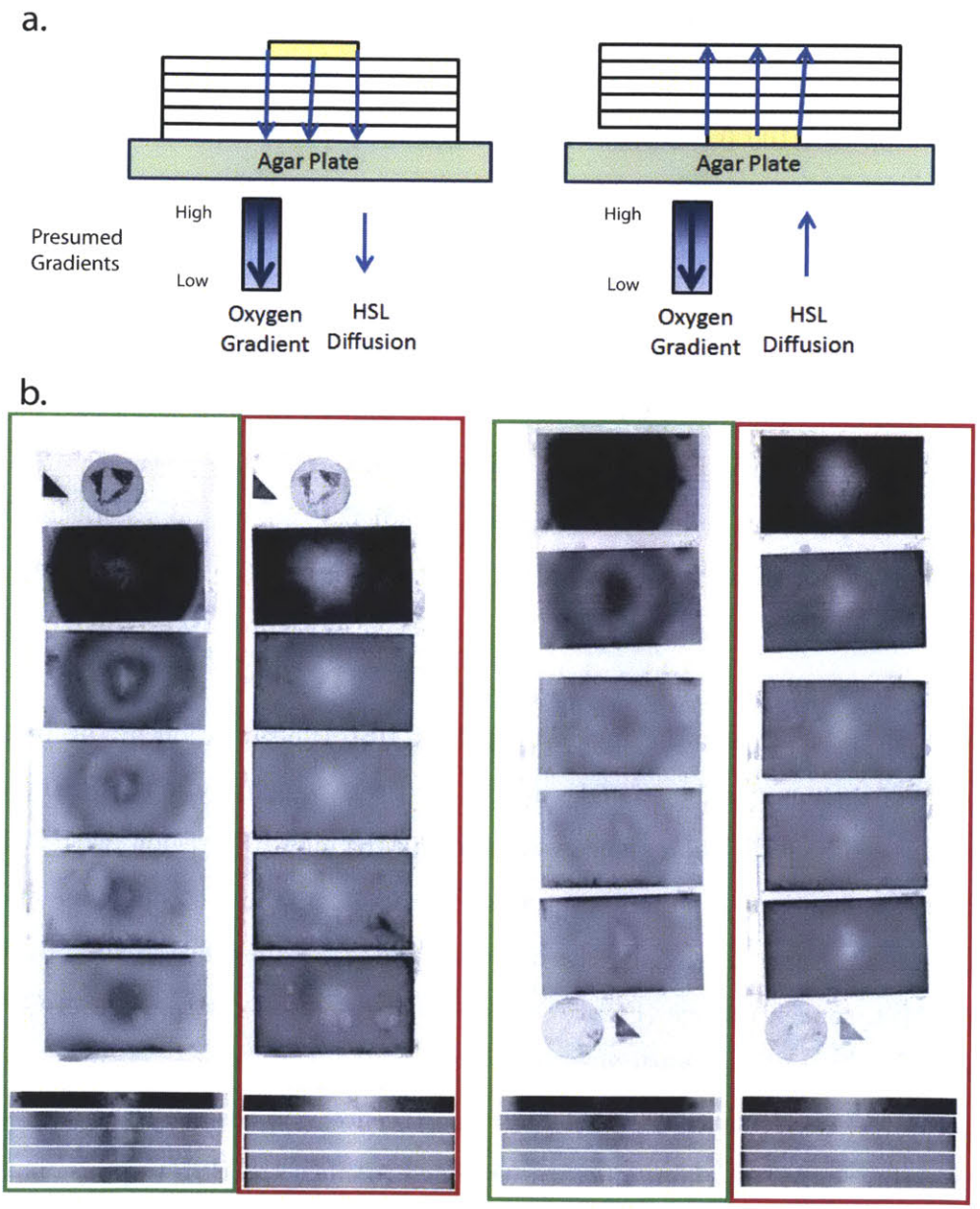


Figure 4-3: Multi-Layer Experiments with "Sender" Located Either above or below "Receiver" *(a)* Schematic of the two experiments where a "Sender" (S strain) containing triangle was put either above (first example) or below (second example) a 5-layer "Receiver" (R strain) stack. A presumed oxygen gradient and direction of HSL diffusion also shown. The supposed O_2 gradient is examined in the Discussion. *(b)* Top and side views of GFP and RFP expression, denoted by green and red rectangles, respectively.

4.4.4 Combination of HSL and Vertical Fluorescence Gradient in Multi-Layer Co-Cultures

The HSL Gradient

Once the ability of the paper-platform to allow for intercellular signaling in co-cultures was determined and the distribution of signal in 2D established, we wanted to characterize chemical signaling-molecule gradients in 3D using the same model organisms. The 3D structure consisted of five stacked rectangles containing the "Receiver" strain and a single triangle containing the "Sender" strain, with a 0.22 μm pore size membrane separating the two strains. The cultures were stacked on LB agar plates with the "Sender" strain-containing triangle either above or below the five layers of "Receiver" strain-containing rectangles. The samples were then incubated at 37°C for 12 h (Figure 4-3 on the facing page a). Stacks were made with the "Sender"-on-top and "Sender"-on-bottom configurations, in order to examine patterns of fluorescence radiating outwards from the "Sender" triangles. We expected to see less HSL response on layers furthest away from the "Sender" strain-containing triangle, similar to the response gradient that we had seen in the 2D binary co-cultures.

The RFP and GFP responses (Figure 4-3 on the preceding page b) formed a ring-pattern similar to that found in the binary co-cultures, in layers close to the "Sender" strain-containing triangle on each of the five "Receiver" containing layers, such as GFP fluorescence in the area under the triangle under the triangle on the layer furthest away from the triangle (Figure 4-3 on the facing page b), the outer radius of the GFP response and the area of no RFP fluorescence that occurred directly above or beneath the sender triangle present throughout all of the layers. The variation in signal response between different layers was not observed to the extent expected. The observation that there was less variation of fluorescent response between layers than expected may be attributed the fact that the five layer stack was only <1 mm thick in total, whereas the rings of GFP fluorescence are close to 4 cm in diameter. Thus, there was more space for signal to move horizontally through one layer than vertically through the different layers. A greater degree of variation may have been observed

had we used a taller stack. Regardless, variation in the fluorescence pattern illustrates the paper-platforms ability to enable visualization of complicated 3D distributions of cellular responses.

The Vertical Fluorescence Gradient

Another unanticipated result of this multi-layer experiment (Figure 4-3 on page 128) was the relative levels of fluorescence based on the layer of the platform and the proximity of the layer to an oxygenated environment. Much greater fluorescence, of both RFP and GFP, was seen on the top layer and sides of the rectangles than in the inner layers. This fluorescence distribution hints at a characteristic found in biofilms, a major type of 3D multicellular structure found in nature in which an oxygen gradient allows aerobic bacteria to be most present near the biofilm surface. This exposed layer of aerobic bacteria also consumes the majority of oxygen present in the structure [328]. The experimental result was further examined by doing a series of experiments in order to determine the number of cells in each layer without the added factor of HSL signaling.

Fluorescence Intensity Calibration Curves

By removing HSL signaling from the system, we sought to investigate this fluorescence distribution in order to determine its cause. Solutions of known "Receiver" concentrations were obtained by centrifuging known volumes of 0.6 OD_{600} "Receiver" strain (10^8) cfu/ml) that had previously been incubated at 37°C, taking the pellets that had formed, and re-suspending them in smaller volumes of LB. Paper squares were then seeded with different concentrations of "Receiver" strain and scanned for fluorescence (Figure 4-4 on the facing page a). This experiment seemed to show a correlation between RFP fluorescence and concentration of "Receiver" on a monolayer.

Cell Counts in Paper

As shown in Figures 4-4 on the next page b and c, five rectangles, inoculated with a 1:100 solution of LB containing "Receiver" bacteria with an optical density of 0.6

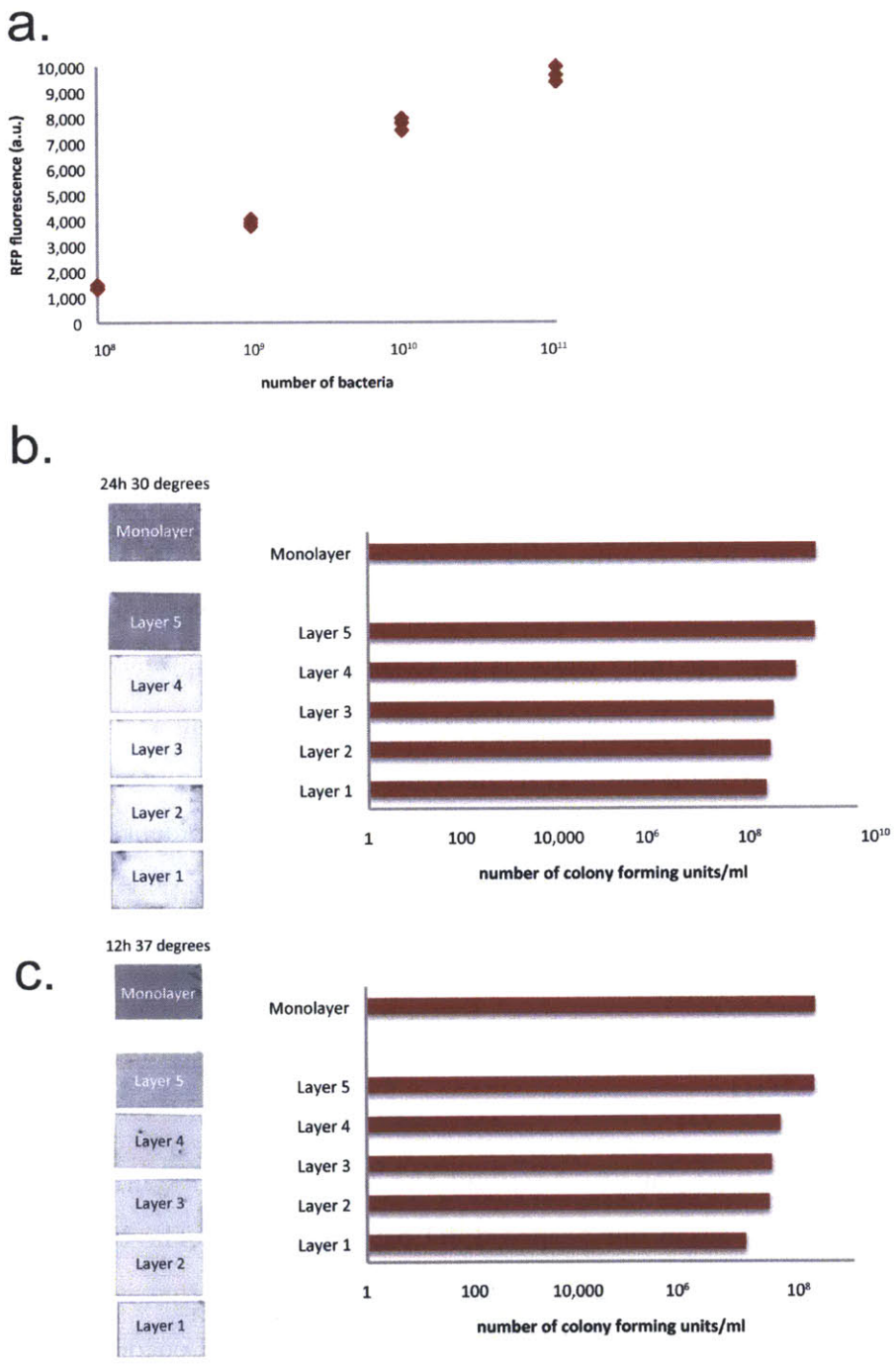


Figure 4-4: RFP Fluorescence Emitted by the "Receiver" in the Absence of HSL and Additional Experiments Done to Characterize RFP Fluorescence by the "Receiver" in the Absence of HSL Signaling *(a)* RFP fluorescence in the "Receiver" (R) strain as a function of the number of bacteria in 1 mL of solution. *(b)* Cell counts per layer of a 5-layer stack and monolayer. The "Receiver" (R) cells were grown in paper at 30°C for 24 h. *(c)* Cell counts per layer of a 5-layer stack and monoculture. The cells were grown in paper at 37°C for 12 h.

OD_{600} (this 1:100 dilution was previously established to be 10^6 cfu/mL), were stacked without an HSL-producing triangle and incubated at 30°C for 24 h and at 37°C for 12 h. The layers were then destacked and scanned for fluorescence. A similar gradient of fluorescence was observed at both temperatures. Each of the layers was then placed in 5 mL and vortexed. "Receiver" strain-containing solutions were obtained. Next, a relative cfu/ml of each layer was determined by diluting and then plating the solution. After incubation, the colonies that formed were then counted (Figures 4-4 on the preceding page b and c). The cell counts may not be absolute, because many cells may have remained in the paper rectangles after vortexing. However, since each rectangle was treated identically, relative cell counts can be compared. These results showed that, at 30°C (Figure 4-4 on the previous page b) there were 10x more cells on the top layer of the multi-layer stack than on the bottom layer (1.4x 10^9 cfu/ml vs. 1.4 x 10^8 cfu/ml,) while at 37°C there were 19x more cells on the top layer when compared to the bottom (1.9 x 10^8 cfu/mL vs. 1.0 x 10^7 cfu/mL). The smaller number of cells on lower layers most likely plays a substantial role in the observed fluorescence. However, other factors such as oxygen-dependent fluorophore maturation will be further examined in the Discussion and may also play roles in this behavior.

4.5 Discussion

A range of experiments were done to determine the behavior of our synthetic *E. coli* intercellular signaling model in our proposed 3D paper platform. The three strains of *E. coli* used were: a "Sender" that excretes HSL; a "Receiver" that is engineered to produce GFP in the presence of HSL along with having RFP encoded under a constitutive promoter; and a "RFP" strain, that has RFP encoded under the same promoter as the "Receiver" (Figure 4-1 on page 120 a). First, we characterized the *E. coli* model. We compared growth rates via fluorescence of the "Receiver" strain grown in paper and directly on agar (Figure 4-1 on page 120 c), the results showed that bacteria grown in both setups reached mid-log phase at essentially the same

time, suggesting that the paper platform did not perturb the *E. coli* growth kinetics. We observed that the paper platform maintained a moist environment surrounding the cells, whereas the LB medium surrounding the spotted cells evaporated after application of the cells to the agar. The moist paper may replicate the environment created by extracellular polymeric substrates that encapsulate cells in biofilms.

We next characterized the fluorescent response of the "Receiver" as a function of HSL concentration at two temperatures, concluding that the RFP fluorescence decreases with increasing HSL concentration (Figure 4-1 on page 120 e), an effect not found in the RFP strain control (Figure 4-1 on page 120 d). GFP fluorescence in the "Receiver" was shown to increase with increasing HSL concentration before declining at high HSL concentrations (Figure 4-1 on page 120 f). Modularity is an important principle in synthetic biology design. To create a modular system we would like to compose two different systems, while preserving their original functions. This experiment highlights that functional composition of modules only occurs under specific conditions. Recently it has become possible to tune biological circuits in bacteria using RBS tuning [264, 302, 352]. An important development is a calculator that allows precise control over downstream protein expression levels [302]) as well as modular BioScaffold parts that allow components to be swapped into a biological circuit at any time after it is created using BioBrick assembly [311]. Prior experimental work has also demonstrated that tuning multiple RBSs within a multi-component biological circuit can lead to large-scale (i.e., greater than 400X) improvements in biological function [352] . This work provides a metric for establishing the biological load of a system if the functional components can be linked to fluorescent protein expression. This work also suggests several routes both to determine and establish system modularity. Growth at lower temperature appears to increase the range of HSL concentrations over which the "Receiver" strain displays modular functional composition of constitutive "RFP expression" and "GFP expression in response to HSL concentration" (i.e., the range where "RFP expression" is constant and "GFP expression" increases with increasing HSL concentration) (see Figure 4-1 on page 120). We could potentially increase this range of modular "Receiver" response by decreasing

the amount of GFP produced relative to a specific HSL concentration in the "Receiver", by tuning the RBS upstream of GFP protein coding region using the RBS calculator [302]). Once the behavior of the *E. coli* "Receiver" was investigated, we could move on to examining intercellular signaling within our platform.

One constraint of our signaling model may be the use of fluorescent proteins as markers for activity. Although it was determined that there was a difference in the number of cells between the top and bottom layers of the 5 layer stacks (Figure 4-4 on page 131), fluorescent protein maturation may also play a factor in the variation of fluorescence between layers. It has been shown that 3D bacterial cultures in nature, such as biofilms, have an oxygen gradient that causes the inner layers of bacteria to be less active [328]. Though the oxygen gradient may explain why there are fewer cells in lower layers, it also may affect fluorescent protein maturation, which uses oxygen as a reactant in its maturation process [61]. Finally, it should be noted that HSL may have peripheral effects on our *E. coli* model that we have not yet investigated. It has been shown in previous studies that HSL may be involved in an *E. coli* pathway involving cell-starvation [164]. However, this potential HSL response does not seem to have a noticeable effect on our signal response system, given that HSL does not decrease fluorescence, even in high concentrations, in the RFP strain (Figure 4-1 on page 120 d).

We used our paper platform to investigate intercellular signaling between our synthetic S and R *E. coli* strains. We first looked at signaling in 2D by producing binary co-cultures with a "Sender" triangle and a "Receiver" rectangle (Figure 4-2 on page 126 b). At the higher temperature culture condition, we identified a larger radius of GFP fluorescence. The binary co-culture system provides visual cues as to the ranges of HSL concentration where the modularity of the system breaks down (Figure 4-2 on page 126 c), and which indicates that incubation at the lower temperature creates a larger range where modular behavior occurs. The non-modular response created an unusual fluorescence pattern when the paper platform was incubated at higher temperatures. We then looked at a 3D signaling pattern by stacking 5 layers of "Receiver" strain containing rectangles above or below a "Sender" strain contain-

ing triangle and determined that variation in HSL response between the layers was present, but did not vary to a great extent (Figure 4-3 on page 128 b). A variation in overall fluorescence between layers, however, was noticed, with more fluorescence in the top layers and on the edges of the sheets of the lower layers. This variation in fluorescence was investigated by doing cell counts of each layer in a stack of R rectangles, which led to the conclusion that significantly fewer cells are present in the lower layers of a stacked platform than are present in the top layer of the platform (Figures 4-4 on page 131 b and c). We expect that a method that allowed increased oxygen diffusion throughout the layers, would increase the uniformity of fluorescence and cellular concentration in the different layers. The many experimental procedures used to characterize intercellular signaling in our synthetic co-culture were aided by the versatility of the paper platform. These experiments also demonstrate the utility of a diffusion gradient in evaluating the ranges of desired operation for a genetic circuit.

The data that we have collected from the experiments presented here demonstrate that our paper platform allows for the analysis of a bacterial co-culture in 3D. We were able to isolate a co-cultures behavior both in 2D (Figure 4-2 on page 126) and 3D (Figure 4-3 on page 128) and determined that a 3D cellular response to a small-molecule signal could exist within the confines of the paper-platform. When we came across an unexpected behavior, the observation of less overall fluorescence in lower layers, we were able to remove the variable of HSL from the experiment and look at how just the variable of layer within the platform affected RFP fluorescence (Figure 4-4 on page 131). Because of the platforms stacked nature, we were able to determine that there was an order of magnitude fewer cells present in the lower parts of the 5-layered platform than in the top layer. Due to the simple, but versatile, nature of the 3D paper platform, we were able to visualize, and investigate, intercellular signaling response within our synthetic *E. coli* model.

A central challenge in synthetic biology is to gain the ability to functionally compose different modular components without changing their function. The experiments here demonstrate a multicomponent system that demonstrates functional composi-

tion under some regimes of operation. Modules in this system can be physically composed through cotransformation of multiple plasmids or through paper seeding and stacking of multiple strains. Physically composed strains are isolated using a common interface (i.e., a track-etch membrane) that allows signal transfer but blocks cell movement. In the paper platform, we can choose whether to place a strain in an oxygenated layer (i.e., the top layer of the system) or an anoxic layer (i.e., the lower layer of the system).

In combination with other recent experimental results, we speculate that it may be possible to extend the range over which modules can be functionally composed in this system. These experiments also demonstrate the utility in dividing either a natural or synthetic system into parts in order to explore its function. Here we explored the HSL response of the RFP circuit both individually in the RFP strain and in combination with the HSL responsive circuit in . The separation of the "Sender" and "Receiver" circuits into different strains also allowed us to example a 5-layer stack of paper seeded with "Receiver" strain bacteria in either the presence or absence of an HSL producing signal. By breaking the system into parts, each of these experiments provided a greater understanding of the underlying parameters that underlie the response curve.

4.6 Acknowledgements

Norville, Pennybaker, Derda, Belcher, and Knight developed the experiments. Norville and Pennybaker performed the experiments. Norville, Derda, and Li developed and implemented analysis techniques. Norville, Pennybaker, Derda, Lockett, Whitesides, Belcher, and Knight drafted the manuscript. Nancy Hopkins, Adam Amsterdam, and Karen Pepper provided insightful feedback that aided in the drafting of the manuscript. We would also like to thank Kelsey Cappelle for her helpful suggestions.

The work was supported by a grant from the National Academy Keck Futures Initiative Synthetic Biology Conference Grant (JN and RD) and a KAUST Scholar Graduate Research Fellowship (JN).

Chapter 5

Conclusion

5.1 Summary of the Experiments

To review I performed experiments with components that added modularity to DNA, protein, and cellular systems. The experiments are described in the abstracts of the manuscripts included below.

5.1.1 From Chapter 2: "Introduction of Customized Inserts for Streamlined Assembly and Optimization of BioBrick Synthetic Genetic Circuits"

BioBrick standard biological parts are designed to make biological systems easier to engineer (e.g. assemble, manipulate, and modify). There are over 5,000 parts (now over 7,000) available in the Registry of Standard Biological Parts that can be easily assembled into genetic circuits using a standard assembly technique. The standardization of the assembly technique has allowed for wide distribution to a large number of users – the parts are reusable and interchangeable during the assembly process. The standard assembly process, however, has some limitations. In particular it does not allow for modification of already assembled biological circuits, addition of protein tags to pre-existing BioBrick parts, or addition of non-BioBrick parts to assemblies.

In this paper we describe a simple technique for rapid generation of synthetic biological circuits using introduction of customized inserts. We demonstrate its use in *E. coli* to express GFP at pre-calculated relative levels and to add an N-terminal tag to GFP. The technique uses a new BioBrick part, called a BioScaffold, that can be inserted into cloning vectors and excised from them to leave a gap into which other DNA elements can be placed. The removal of the BioScaffold is performed by a Type IIB restriction enzyme that recognizes the BioScaffold but cuts into the surrounding sequences; therefore, the placement and removal of the BioScaffold allows the creation of seamless connections between arbitrary DNA sequences in cloning vectors. The BioScaffold contains a built-in RFP reporter; successful insertion of the BioScaffold is, thus, accompanied by gain of red fluorescence and its removal is manifested by disappearance of the red fluorescence.

The ability to perform targeted modifications of existing BioBrick circuits with BioScaffolds (1) simplifies and speeds up the iterative design-build-test process through direct reuse of existing circuits, (2) allows incorporation of sequences incompatible with BioBrick assembly into BioBrick circuits (3) removes scar sequences between standard biological parts, and (4) provides a route to adapt synthetic biology innovations to BioBrick assembly through the creation of new parts rather than new assembly standards or parts collections.

5.1.2 From Chapter 3: "Fast and Easy Protocol for the Purification of Recombinant S-Layer Protein for Synthetic Biology Applications"

A goal of synthetic biology is to make biological systems easier to engineer. One of the aims in the engineering of biomaterials is to design, with nanometer-scale precision, surfaces with well-defined properties. The surface-layer protein SbpA forms 2D arrays naturally on the cell surface of *L. sphaericus* but also as purified protein in solution upon addition of divalent cations. Its high propensity to form crystalline arrays, the simple way by which its crystallization can be controlled by divalent cations and

the possibility to genetically alter the protein make SbpA an attractive molecule for the design of surfaces with defined properties. To be a useful tool, however, it is important that a simple protocol can be used to produce recombinant wild-type as well as modified SbpA in large quantities and in a biologically active form. The present study addresses this requirement by introducing a facile method to produce large quantities of recombinant, active SbpA. The purification procedure we introduce is mild and non-denaturing.

5.1.3 From Chapter 3: 7 Å Projection Map of the S-Layer Protein SbpA Obtained with Trehalose-Embedded Monolayer Crystals

2D crystallization on lipid monolayers is a versatile tool to obtain structural information of proteins by electron microscopy. An inherent problem with this approach is to prepare samples in a way that preserves the crystalline order of the protein array and produces specimens that are sufficiently flat for high-resolution data collection at high tilt angles. As a test specimen to optimize the preparation of lipid monolayer crystals for electron microscopy imaging, we used the S-layer protein SbpA, a protein with potential for designing arrays of both biological and inorganic materials with engineered properties for a variety of nanotechnology applications. Sugar embedding is currently considered the best method to prepare 2D crystals of membrane proteins reconstituted into lipid bilayers. We found that using a loop to transfer lipid monolayer crystals to an electron microscopy grid followed by embedding in trehalose and quick-freezing in liquid ethane also yielded the highest resolution images for SbpA lipid monolayer crystals. Using images of specimens prepared in this way we could calculate a projection map of SbpA at 7 Å resolution, one of the highest resolution projection structures obtained with lipid monolayer crystals to date.

5.1.4 From Chapter 4: "Paper Supported Polycultures of Synthetically Programmed Bacteria"

Many types of bacteria have the ability to form a 3D-structure known as a biofilm. Segregated into layers, the bacteria in biofilms are able to communicate with each other, a process that regulates biofilm formation and growth. In synthetic biology, it would be useful to construct systems that have a 3D structure analogous to biofilms, but where we can control the location of individual strains. This platform could be used to study analogues of natural systems such as biofilms or to construct synthetic structures with novel properties. We present here a platform for constructing 3D arrangements of bacteria. The platform contains stacked layers of paper that are seeded with bacteria. Bacteria are grown within the paper structure, and then later destacked and analyzed. Membranes can be used to separate layers that contain separate strains. Visualization of each separate layer provides a 3D representation of the cells growing within the structure. By using two types of engineered bacteria, we show how bacteria within the model can communicate with each other and how bacterial communication can be added, and monitored. We learned that distance-dependent signaling and oxygen gradients, similar to what is found in natural biofilms, could be observed using the platform.

5.2 Analysis

Modular components are found throughout biological systems; however, the naturally modular systems present are not easy to engineer. Biological systems are modular relative to evolution and have characteristics that can be optimized efficiently by evolution. Evolution latches onto systems that afford local changes without breaking the surrounding system. In creating tools for synthetic biology, it is also beneficial to have the ability to change some parts of a system without changing everything, but evolution only optimizes systems through random mutation followed by natural selection. That does not necessarily favor systems where the connection between a

change and the altered effects is obvious, and it does not necessarily create systems that easy to alter in a calculated fashion. Therefore, one approach is to refactor the natural modularity of systems into where specific components of the systems can be changed explicitly.

The BioScaffold part adds engineering control to natural biological systems, in places where changes at specific locations in DNA sequences effect beneficial changes. For example, in bacterial operons, mutations in a promoter can greatly affect the expression level of all downstream proteins in a pathway [?], while maintaining their expression ratios. Alternatively, mutations in the ribosome binding site (RBS) upstream of a specific protein can greatly alter the expression level of the downstream protein [321]. Conversely, mutations in a protein coding region may create silent or minimal changes due to the fault tolerance and redundancy of the genetic code [354]. From an engineering perspective, BioScaffold parts can be used to abstract the ability to alter the concentration of specific enzymes or the expression level of the pathway. Thus, in this system specific changes can be made to alter this system rather than waiting for mutations to alter the system in a positive or negative fashion.

The S-layer protein SbpA brings engineering control to a natural technique for optimizing reactions, by providing a simple way to bring together enzymes in close proximity in defined ratios, while using standard interfaces to mediate the connections between components. The analogous natural systems are bacterial microcompartments, which are protein shells of fixed size that bring together the enzymes for a specific pathway [363]. The pores of the cage are optimized for the input and output of specific chemicals. SbpA provides a standard way to bring enzymes together in close proximity. It has been demonstrated that proteins can be fused to a specific connection point on the C-terminus of SbpA without interfering with the crystallization properties of the protein. SbpA crystallized using specific domains. Thus, this system provides a standard interface that brings enzymes together. The types of enzymes used and their concentrations can be tuned as appropriate.

The cells-on-paper system brings engineering control to the ability of biological systems to spatially pattern and program different cell types. In developmental pro-

cesses for multicellular organisms, Hox genes spatially pattern different cell types and create containers for different downstream processes [273]. In the cells-on-paper system, cell impermeable membranes form the interface between different compartments and paper that contains cells and media occupies the compartment. One can easily mix and match the types or genetic programs of a strain used in a specific compartment and change the locations of the compartments relative to each other. This system provides a route to create defined compartments, while allowing communication along a single interface.

5.3 Impact

"Engineering biological systems requires a fundamentally different viewpoint from the science of biology. Key engineering principles of modularity, simplicity, separation of concerns, abstraction, flexibility, hierarchical design, isolation, and standardization are of critical importance. The essence of engineering is the ability to imagine, design, model, build, and characterize novel systems to achieve specific goals. Current tools and components for these tasks are primitive" [178]. In this dissertation, I have created modular tools for engineering linear DNA sequences, 2D protein structures, and 3D cellular ensembles that other researchers can use. I have also demonstrated how one can identify modular natural biological systems and then refactor them so that they are easier to engineer, build upon, and use as modules in larger systems. The process used to design the modules can also be used to create new engineering tools as well as to adapt existing tools for new functions and chassis organisms.

Appendix A

Abbreviations and Supplementary Information

A.1 Supplementary Information for Examination of Non-Denatured SbpA throughout the Crystallization Process in Chapter 3

The Supplementary Information included in the Sections below is from the manuscript Julie E. Norville, Deborah F. Kelly, Thomas F. Knight, Jr., Angela M. Belcher, and Thomas Walz. "Fast and easy protocol for the purification of recombinant S-layer protein for synthetic biology applications." [265], the main text of which is included in Section 3.2 on page 80 to Section 3.8 on page 90 of Chapter 3.

A.1.1 Supplementary Introduction

Expression of SbpA at 37°C and at Low Temperature

Several groups have already developed and utilized purification protocols for both native and recombinant SbpA variants [39, 87, 166, 191, 257, 263, 332, 336], but in all protocols the protein is denatured at some point during the purification procedure. Purification of native SbpA requires denaturing conditions to separate the S-layer

protein from the cell wall [39, 263]. When expressed in *E. coli*, recombinant SbpA has been found in inclusion bodies that have to be dissolved using denaturing conditions and refolded [257]. Our protocol uses expression at low temperature, which improves the solubility of expressed protein [135, 289]. Inclusion bodies are thought to form when protein is produced so fast that it does not have enough time to fold properly and aggregates [42, 135]. At lower temperatures, transcription and translation slow down and the strength of hydrophobic interactions that contribute to protein misfolding is reduced, thus increasing the chances for the expressed protein to fold into its native conformation [42, 135].

A.1.2 Supplementary Materials and Methods

Purification of Inclusion Bodies

SbpA-His_7 was expressed in *E. coli* HMS174 (DE3) (EMD Biosciences). The culture was grown at 37°C in LB medium containing 50 µg/ml kanamycin and 50 µg/ml rifampicin until the OD_{600} reached 0.6. Upon induction by addition of 0.5 mM IPTG, the incubation temperature was either maintained at 37°C for 5 h after induction [257] or lowered to 18°C for 18 h after induction. The cells were then pelleted by centrifugation at 16,000g for 15 minutes at 4°C. The cell pellet was resuspended in lysis buffer [129] containing 20 mM Tris, pH 8.0, 300 mM NaCl, 10 mM imidazole and 1 mg/ml lysozyme (EMD Biosciences) and sonicated on ice. Unbroken cells were pelleted [86] by centrifugation at 500g for 5 minutes. Inclusion bodies were then pelleted from the soluble fraction of the crude cell lysate by centrifugation at 20,000g for 30 minutes [173].

Monolayer Crystallization of S-Layer Proteins

1.9 µl of native or recombinant SbpA solution (7.6 nmol) was added to 23.1 µl reaction buffer (20 mM Tris, pH 8, 100 mM NaCl, 50 mM $CaCl_2$) and the reaction mixture was placed into a well in a teflon block as described previously [263]. The sample was overlayed with 1 µl of a lipid mixture containing 0.25 µg DDMA and 0.55 µg DMPC

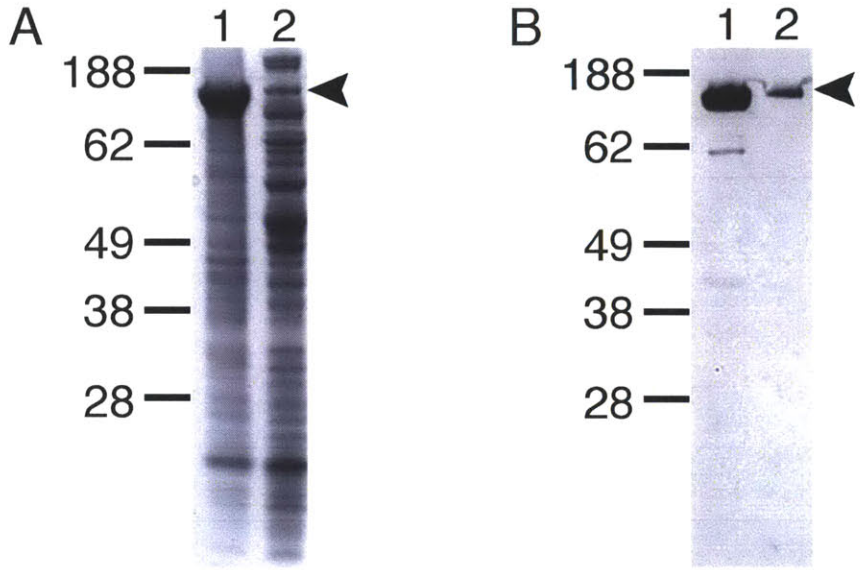


Figure A-1: Analysis of Recombinant SbpA-His$_7$ in the Inclusion Bodies and Soluble Fraction after 5 Hours of Expression [257] at 37°C. The soluble fraction and inclusion bodies were analyzed by SDS-PAGE *(A)* and Western blot analysis with anti-His antibody *(B)*. Lane1, inclusion bodies; Lane 2, soluble fraction. Experiments were reproducible within three different preparations.

in chloroform. After a 24 h incubation at room temperature, the monolayer crystals were transferred to continuous carbon EM grids and negatively stained as described [263].

A.1.3 Supplementary Results and Discussion

Partitioning of SbpA-His$_7$ between the Inclusion Bodies and the Soluble Fraction, While Being Expressed at 37°C or 18°C

To assess whether all SbpA-His$_7$ was expressed exclusively in soluble form at 18°C, we examined *E. coli* cells for inclusion bodies. Compared to cells grown at 37°C, in which 83% of the protein localized to inclusion bodies (Figure A-1), only 15% of the protein was in inclusion bodies when expression was performed at 18°C (Figure A-2 on the following page). Our purification protocol yields about 40 mg of soluble SbpA from 5 g of wet mass, whereas denaturing purification of recombinant SbpA yields

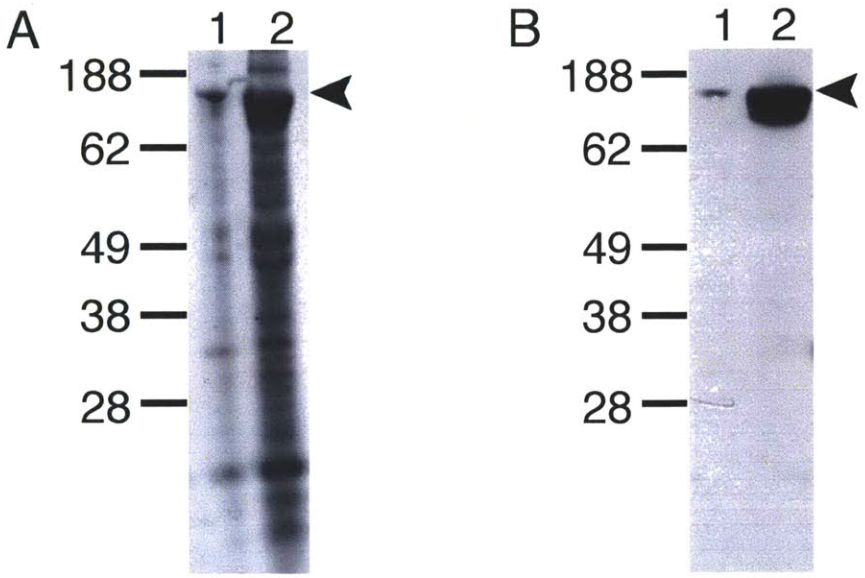


Figure A-2: Analysis of Recombinant SbpA-His$_7$ in the Inclusion Bodies and Soluble Fraction after 18 Hours of Expression at 18°C. The soluble fraction and inclusion bodies were analyzed by SDS-PAGE *(A)* and Western blot analysis with anti-His antibody *(B)*. Lane1, inclusion bodies; Lane 2, soluble fraction. Experiments were reproducible within three different preparations.

120 mg [257] and denaturing purification [263] of native SbpA yields 60 mg. Further details on the expression and purification of SbpA are provided in Supplementary Methods (i.e., Subsection A.1.2 on page 144).

Monolayer Crystallization of Recombinant SbpA

SbpA usually forms crystalline arrays on the bacterial cell surface. To verify that recombinant SbpA proteins retain the ability to form 2D crystals on surfaces, we used native SbpA as well as the recombinant proteins to form crystals on lipid monolayers as described before [263]. For each sample, three specimens were prepared and examined by negative stain EM. The recombinant SbpA-His$_7$ (Figure A-3 on the facing page A) and SbpA$_{C-trunc}$-His$_7$ (Figure A-3 on the next page B) proteins formed crystals as efficiently as native SbpA (Figure A-3 on the facing page C). For these proteins, 5-8 crystals of size 0.5 to 2 μm were present on 80% of the grid squares.

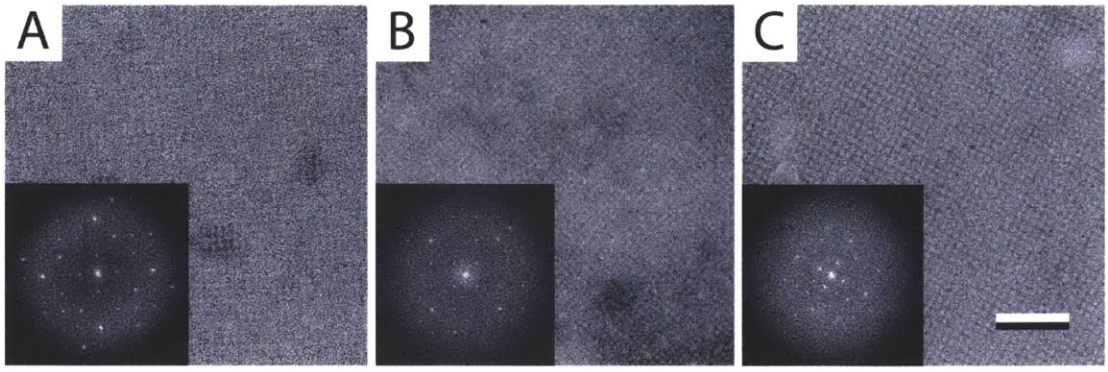


Figure A-3: Crystalline Arrays Formed by Native and Recombinant SbpA on Lipid Monolayers. *(A)* SbpA-His_7, *(B)* $SbpA_{C-trunc}$-His_7, and *(C)* native SbpA. Scale bar is 100 nm. SbpA-His_7 and native SbpA experiments were reproducible within ten different preparations. $SbpA_{C-trunc}$-His_7 experiments were reproducible within three different preparations.

However, the N-terminally truncated $SbpA_{N-trunc}$-His_7 protein consistently failed to form crystals, indicating that the missing N-terminal domain is required for crystallization on the lipid monolayer. The lattice constants of the arrays formed by the various SbpA variants were essentially identical to those of arrays formed in solution, namely a = b = 131 ± 0.6 Å for SbpA-His_7, a = b = 131 ± 1 Å for $SbpA_{C-trunc}$-His_7, and a = b = 131 ± 1 Å for the native protein.

Appendix B

Permission to Use Previously Published Manuscripts

B.1 Chapter 2: The Manuscript Entitled "Introduction of customized inserts for streamlined assembly and optimization of BioBrick synthetic genetic circuits"

Permission is granted for inclusion of the manuscript [264] through the BioMed Central Open Access license agreement, which is quoted below:

BioMed Central Open Access license agreement
Brief summary of the agreement

Anyone is free:

-to copy, distribute, and display the work;
-to make derivative works;
-to make commercial use of the work;

Under the following conditions: Attribution

-the original author must be given credit;
-for any reuse or distribution, it must be made clear to others what the license terms of this work are;
-any of these conditions can be waived if the author gives permission.

Statutory fair use and other rights are in no way affected by the above.

B.2 Chapter 3: The Manuscript Entitled "Fast and Easy Protocol for the Purification of Recombinant S-Layer Protein for Synthetic Biology Applications"

Permission is granted for inclusion of the manuscript [263] in the dissertation based on the permitted uses in the Wiley-Blackwell Copyright Transfer Agreement. The manuscript is less than half the material in the dissertation. Only minor formatting changes (i.e., bibliography, citations, layout) have been made in order to integrate the manuscript into the dissertation. The relevant section (C., 3. b.) is quoted below:

C. PERMITTED USES BY CONTRIBUTOR

3. Final Published Version. Wiley-Blackwell hereby licenses back to the Contributor the following rights with respect to the final published version of the Contribution:

b. Re-use in other publications.The right to re-use the final Contribution or parts thereof for any publication authored or

edited by the Contributor (excluding journal articles) where such re-used material constitutes less than half of the total material in such publication. In such case, any modifications should be accurately noted.

B.3 Chapter 3: The Manuscript Entitled"7 Å Projection Map of the S-Layer Protein SbpA Obtained with Trehalose-Embedded Monolayer Crystals"

Permission is granted for inclusion of the manuscript [263] through the Authors' Rights Responsibilities as set by Elsevier on their website and quoted below:

What rights do I retain as a journal author*?
-the right to include the journal article, in full or in part, in a thesis or dissertation;

This information is presented more formally in a guide distributed by the publisher [8] (http://libraryconnect.elsevier.com/lcp/0403/lcp0403.pdf).

HOW AUTHORS CAN REUSE THEIR OWN ARTICLES
PUBLISHED BY ELSEVIER

General use of articles

Authors publishing in Elsevier journals retain wide rights to continue to use their works to support scientific advancement, teaching and scholarly communication.

An author can, without asking permission, do the following after publication of the authors article

in an Elsevier-published journal:

-Include the article in full or in part in a thesis or dissertation.

Bibliography

[1] BBF Standards/Technical/Formats. [http://bbf.openwetware.org/RFC.html], (Archived by WebCite at http://www.webcitation.org/5u3WYySIC). 57

[2] BBa_J04430 GFP reporter circuit. [http://partsregistry.org/Part:BBa$_$J04430], (Archived by WebCite at http://www.webcitation.org/5u3ddROiU). 69

[3] International Genetically Engineered Machine Competition (iGEM). [http://ung.igem.org/About], (Archived by WebCite at http://www.webcitation.org/5u3ZrbtLO). 59

[4] Registry of Standard Biological Parts. [http://partsregistry.org], (Archived by WebCite at http://www.webcitation.org/5u3dGQB0G). 58, 71

[5] *Marine engineering and shipping review.* Number 42. Simmons Boardman Pub. Co., 1937. 22

[6] Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations. *Eur. J. Biochem*, 150:1–5, 1984. 59, 62

[7] *Size limits of very small microorganisms: proceedings of a workshop.* Compass series. National Academy Press, 1999. 45

[8] *Ways to use journal articles published by Elsevier: A practical guide.* Pamphlet 4. Elsevier, 3 edition, 2008. 151

[9] Restriction enzyme specifications. [http://www.sibenzyme.com], 2009.

[10] Hammer News. [http://www.hammar.com.au/pdf/15_years_in_australia.pdf], 2010. 25

[11] Butterfly picture. 2011. URL http://welovebutterflies.com/wp-content/uploads/2010/10/butterfly-pictures-1.jpg. 22

[12] Clontech FAQs. [http://www.clontech.com/products/detail.asp?tabno=2&catalog$_$id=632481&page=all&faq$_$id=154555], (Archived by WebCite at http://www.webcitation.org/5u3ZzRkb2), 2011. 68, 69

[13] CMA CGM group container fleet. [http://www.anl.com.au/docs/guides/Container_Fleet_Guide.pdf], 2011. 24

[14] Garmentainer specification. [http://www.oocl.com/eng/ourservices/containers/containerspecifications/Pages/garmentainers.aspx], 2011. 24

[15] The RBS calculator. [http://www.voigtlab.ucsf.edu/software/], (Archived by WebCite at http://www.webcitation.org/5u4s1cfaq), 2011. 73

[16] The RBS calculator FAQ. [https://salis.psu.edu/software/static/faq.html], (Archived by WebCite at http://www.webcitation.org/5u3aFeee3), 2011. 73

[17] The RBS calculator FAQ. [http://voigtlab.ucsf.edu/software/faq.html], (Archived by WebCite at http://www.webcitation.org/5u3aRZplP), 2011. 73

[18] Refrigerated container specification. [http://www.apl.com/equipment/html/equipment_specs_refrig.html], 2011. 24

[19] 1945 VW Assembly Line. 2011. URL http://www.shorey.net/auto/german/volkswagon/beetle/htmltree.html. 22

[20] Protocol for BioBrick™ Assembly Kit. [http://www.neb.com/nebecomm/products/protocolProductE0546.asp], 2012. 119

[21] A. Abzhanov, M. Protas, B. R. Grant, P. R. Grant, and C. J. Tabin. *Bmp4* and morphological variation of beaks in Darwin's finches. *Science*, 305:1462–1465, 2004. 19, 42

[22] A. Abzhanov, W. P. Kuo, C. Hartmann, B. R. Grant, P. R. Grant, and C. J. Tabin. The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature*, 442:563–567, 2006. 42, 43

[23] R. K. G. adn R. R. Strathmann. The evolution of multicellularity: a minor major transition? *Annu. Rev. Ecol. Evol. Syst.*, 38:621–654, 2007. 32

[24] U. Aebi, P. R. Smith, J. Dubochet, C. Henry, and E. Kellenberger. A study of the structure of the T-layer of *Bacillus brevis*. *J. Supramol. Str. Cell*, 1: 498–522, 1974. 111

[25] B. Afonso, P. A. Silver, and C. M. Ajo-Franklin. A synthetic circuit for selectively arresting daughter cells to create aging populations. *Nucleic Acids Res.*, 38:2727–2735, 2010. 59

[26] C. M. Agapakis, D. C. Ducat, P. M. Boyle, E. H. Wintermute, J. C. Way, and P. A. Silver. Insulation of a synthetic hydrogen metabolism circuit in bacteria. *J. Biol. Eng.*, 4:3, 2010. 59

[27] H. Agerschou. *Planning and design of ports and marine terminals.* Thomas Telford, 2004. 22

[28] J. Aleksic, A. Bizzarri, Y. Cai, B. Davidson, K. de Mora, S. Ivakhno, S. L. Seshasayee, J. Nicholson, J. Wilson, A. Elfick, C. French, L. Kozma-Bognar,

H. Ma, and A. Millar. Development of a novel biosensor for the detection of arsenic in drinking water. *IET Synthetic Biology*, 1:87–90, 2007. 59

[29] H. Alper, C. Fischer, E. Nevoigt, and G. Stephanopoulas. Tuning genetic control through promoter engineering. *Proc. Natl. Acad. Sci. U.S.A.*, 102:12678–12683, 2004. 76

[30] S. Aluru. *Handbook of computational molecular biology.* Chapman & Hall/CRC computer and information science series. Chapman & Hall/CRC, 2006. 45

[31] L. A. Amos, F. van den Ent, and J. Lowe. Structural/functional homology between the bacterial and eukaryotic cytoskeletons. *Curr. Opin. Cell Biol.*, 16: 24–31, 2004. 30

[32] J. C. Anderson, J. E. Dueber, M. Leguia, G. C. Wu, J. A. Goler, A. P. Arkin, and J. D. Keasling. BglBricks: A flexible standard for biological part assembly. *J. Biol. Eng.*, 4:1, 2010. 57, 58

[33] J. C. Anderson, J. E. Dueber, M. Leguia, G. C. Wu, J. A. Goler, A. P. Arkin, and J. D. Keasling. BBF 21: BglBricks assembly standard. [`http://dspace.mit.edu/handle/1721.1/46747`], (Archived by WebCite at `http://www.webcitation.org/5u3Xfg3Gd`), *DSPACE*, 2009. 47, 57

[34] C. B. Anfinsen. Principles that govern the folding of protein chains. *Science*, 181:223–230, 1973. 46

[35] A. P. Arkin. Synthetic cell biology. *Curr. Opin. Biotech.*, 12:638–644, 2001. 58

[36] F. J. Asturias and R. D. Kornberg. A novel method for transfer of two-dimensional crystals from the air/water interface to specimen grids: EM sample preparation/lipid-layer crystallization. *J. Struct. Biol.*, 114:60–66, 1995. 93, 96, 110

[37] A. J. Avila-Sakar and W. Chiu. Visualization of beta-sheets and side-chain clusters in two-dimensional periodic arrays of streptavidin on phospholipid mono-

layers by electron crystallography. *Biophys. J.*, 70:57–68, 1996. 92, 93, 109, 110, 111, 112

[38] S. Ayukawa, A. Kobayashi, Y. Nakashima, H. Takagi, S. Hamada, M. Uchiyama, K. Yugi, S. Murata, Y. Sakakibara, M. Hagiya, M. Yamamura, and D. Kiga. SYNAC: SYnthetic biological Automation for Noughts and Crosses. *IET Synthetic Biology*, 1:64–67, 2007. 59

[39] B. Schuster and E. Györvary and D. Pum and U. Sleytr. *Protein Nanotechnology: Protocols, Instrumentation, and Applications*, chapter Nanotechnology with S-layer proteins, pages 101–124. Humana Press, 2005. 82, 94, 143, 144

[40] D. Baker and A. Sali. Protein structure prediction and structural genomics. *Science*, 294:93–96, 2001. 45

[41] C. Y. Baldwin and K. B. Clark. *Modularity in the design of complex engineering systems.* Complex Engineered Systems: Science Meets Technology, New England Complex Systems Institute Series on Complexity. Springer–Verlag, New York, NY, 2006. 21

[42] F. Baneyx and M. Mujacic. Recombinant protein folding and misfolding in *Escherichia coli*. *Nat. Biotechnol.*, 22:1399–1408, 2004. 144

[43] K. Bansal, K. Yang, G. B. Nistala, R. B. Gennis, and K. D. Bhalerao. A positive feedback-based gene circuit to increase the production of a membrane protein. *J. Biol. Eng.*, 4:6, 2010. 57

[44] C. J. Bashor, N. C. Helman, S. Yan, and W. A. Lim. Engineering synthetic signaling proteins with ultrasensitive input/output control. *Nat. Biotechnol.*, 25:660–662, 2007. 47

[45] C. J. Bashor, N. C. Helman, S. Yan, and W. A. Lim. Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. *Science*, 319: 1539–1543, 2008. 47

[46] B. L. Bassler. Small talk: cell-to-cell communication in bacteria. *Cell*, 109: 421–424, 2002. 116

[47] S. Basu, Y. Gerchman, C. H. Collins, F. H. Arnold, and R. Weiss. A synthetic multicellular system for programmed pattern formation. *Nature*, 434:1130–1134, 2005. 48, 117, 124

[48] W. Baumeister, I. Wildhaber, and H. Engelhardt. Bacterial surface proteins. Some structural, functional and evolutionary aspects. *Biophys. Chem.*, 29:39–49, 1988. 94

[49] J. Baumgardner, K. Acker, O. Adefuye, S. T. Crowley, W. DeLoache, J. O. Dickson, L. Heard, A. T. Martens, N. Morton, M. Ritter, A. Shoecraft, J. Treece, M. Unzicker, A. Valencia, M. Waters, A. M. Campbell, L. J. Heyer, J. L. Poet, and T. T. Eckdahl. Solving a Hamiltonian Path Problem with a bacterial computer. *J. Biol. Eng.*, 3:11, 2009. 59

[50] A. Beeby and A. Brennan. *First ecology: ecological principles and environmental issues.* Oxford University Press, 2008. 30

[51] L. Bejder and B. K. Hall. Limbs in whales and limblessness in other vertebrates: mechanisms of evolutionary and developmental transformation and loss. *Evo. Devo.*, 4:445–458, 2002. 41

[52] M. Bencina and R. Jerala. BBF RFC 37: Fusion protein BioBrick assembly standard with optional linker extension. [`http://dspace.mit.edu/handle/1721.1/46705`], (Archived by WebCite at `http://www.webcitation.org/5u3YpNbqX`), *DSPACE*, 2009. 57

[53] S. A. Benner and A. M. Sismour. Synthetic biology. *Nat. Rev. Genet.*, 6: 533–543, 2005. 45

[54] D. H. Bergey and J. G. Holt. *Bergey's manual of determinative bacteriology.* Williams & Wilkins, 1994. 31

[55] K. Bernhardt, E. J. Carter, N. S. Chand, J. Lee, Y. Xu, X. Zhu, J. W. Ajioka, J. M. Goncalves, J. Haseloff, G. Micklem, and D. Rowe. New tools for self-organized pattern formation. *IET Synthetic Biology*, 1:29–31, 2007. 59

[56] K. D. Bhalerao. Synthetic gene networks: the next wave in biotechnology? *Trends. Biotechnol.*, 27:368–374, 2009. 57

[57] W. H. Bingle, J. F. Nomellini, and J. Smit. The secretion signal of the *Caulobacter crescentus* S-layer protein is located within the C-terminal 82 amino acids of the molecule. *J. Bacteriol.*, 182:3298–3301, 2000. 81

[58] J. Black. *Microbiology: principles and explorations.* J. Wiley & Sons, 2008. 32

[59] J. S. Bonifacino and B. S. Glick. The mechanisms of vesicle budding and fusion. *Cell*, 116:153–166, 2004. 36, 37

[60] M. E. Boraas, D. B. Seale, and J. E. Boxhorn. Phagotrophy by a flagellate selects for colonial prey: a possible origin of multicellularity. *Evol. Ecol.*, 12: 153–164, 1998. 38

[61] G. Borriello, E. Werner, F. Roe, A. M. Kim, G. D. Ehrlich, and P. S. Stewart. Oxygen limitation contributes to antibiotic tolerance of *Pseudomonas aeruginosa* in biofilms. *Antimicrob. Agents Chemother.*, 48:2659–2664, 2004. 134

[62] J. Böse. *Handbook of Terminal Planning.* Operations Research Computer Science Interfaces. Springer, 2011. 23

[63] J. U. Bowie, J. F. Reidhaar-Olson, W. A. Lim, and R. A. Sauer. Deciphering the message in protein sequences: tolerance to amino acid substitutions. *Science*, 247:1306–1310, 1990. 47

[64] R. A. Bradshaw. Protein translocation and turnover in eukaryotic cells. *TIBS*, 14:276–279, 1989. 37

[65] K. Brenner, D. K. Karig, R. Weiss, and F. A. Arnold. Engineered bidirectional communication mediates a consensus in a microbial biofilm consortium. *Proc. Natl. Acad. Sci. U.S.A.*, 104:17300–17304, 2007. 75, 76

[66] A. Brewer, K. Button, and D. Hensher. *Handbook of logistics and supply-chain management.* Handbooks in transport. Pergamon, 2001. 25

[67] A. Brisson, W. Bergsma-Schutter, F. Oling, O. Lambert, and I. Reviakine. Two-dimensional crystallization of proteins on lipid monolayers at the air-water interface and transfer to an electron microscopy grid. *J. Cryst. Growth*, 196: 456–470, 1999. 93

[68] R. R. Burgess and M. W. Knuth. *Strategies for Protein Purification and Characterization: A Laboratory Course Manual*, chapter Purification of a recombinant protein overproduced in *Escherichia coli.* Cold Spring Harbor Laboratory Press, 1996. 83

[69] A. C. Burke, C. E. Nelson, B. A. Morgan, and C. Tabin. *Hox* genes and the evolution of vertebrate axial morphology. *Development*, 121:333–346, 1995. 39

[70] D. R. Burrill and P. A. Silver. Making cellular memories. *Cell*, 140:13–18, 2010. 59

[71] C. Huber and N. Ilk and D. Rünzler and E. M. Egelseer and S. Weigert and U. B. Sleytr and Margit Sara. The three S-layer-like homology motifs of the S-layer protein SbpA of *Bacillus sphaericus* CCM 2177 are not sufficient for binding to the pyruvylated secondary cell wall polymer. *Mol. Microbiol.*, 55: 197–205, 2004. 82, 87, 89

[72] C. Kaas and N. B. Hansen and H. J. Genee and L. R. Olsen and C. Matos and M. T. Bonde. BBF RFC 39: The USER cloning standard. [`http://dspace.mit.edu/handle/1721.1/49522`], (Archived by WebCite at `http://www.webcitation.org/5u3ZOGD6G`), *DSPACE*, 2009. 57

[73] C. Marley. Aesthetica sphere. Accessed in Sept 2011. URL http://goldbugpasadena.com/christopher-marley/christopher-marley-aesthetica-sphere-in-20x24-inch-black-frame. 20

[74] C. Völlenkle and S. Weigert and N. Ilk and E. Egelseer and V. Weber and F. Loth and D. Falkenhagen and U. B. Sleytr and M. Sara. Construction of a functional S-layer fusion protein comprising an immunoglobulin G-binding domain for development of specific absorbents for extracorporeal blood purification. *Appl. Environ. Microbiol.*, 70:1514–1521, 2004. 89

[75] B. Calcott and K. Sterelny. *The major transitions in evolution revisited.* MIT Press, 2011. 31, 39, 41

[76] B. Canton, A. Labno, and D. Endy. Refinement and standardization of synthetic biological parts and devices. *Nat. Biotechnol.*, 26:787–793, 2008. 59

[77] S. D. Carlo, M. Adrian, P. Kalin, J. M. Mayer, and J. Dubochet. Unexpected property of trehalose as observed by cryo-electron microscopy. *J. Microsc.*, 196: 40–45, 1999. 109

[78] S. B. Carroll. Homeotic genes and the evolution of arthropods and chordates. *Nature*, 10:479–485, 1995.

[79] S. B. Carroll. Chance and necessity: the evolution of morphological complexity and diversity. *Nature*, 409:1102–1109, 2001. 38

[80] J. P. Caviston and E. L. F. Holzbaur. Microtubule motors at the intersection of trafficking and transport. *Trends Cell. Biol.*, 18:530–537, 2006. 37

[81] H. Celia, E. Wilson-Kubalek, R. A. Milligan, and L. Teyton. Structure and function of a membrane-bound murine MHC class I molecule. *Proc. Natl. Acad. Sci. U.S.A.*, 196:456–470, 1999. 92

[82] M. Chalfie and S. R. Kain. *Green fluorescent protein: properties, applications, and protocols.* Wiley–Liss, New York, NY, 1998. 68, 69

[83] A. Che. BioBricks++: Simplifying assembly of standard DNA components. [http://dspace.mit.edu/handle/1721.1/39832], (Archived by WebCite at http://www.webcitation.org/5u3YJEYj9), *DSPACE*, 2004. 57

[84] A. Che. BBF RFC 20: Constraint relaxation of RFC 10 for assembling standard biological parts. [http://dspace.mit.edu/handle/1721.1/44962], (Archived by WebCite at http://www.webcitation.org/5u3Y5MqXU), *DSPACE*, 2009. 47, 57

[85] I.-G. Choi and S.-H. Kim. Evolution of protein structural classes and protein sequence families. *Proc. Natl. Acad. Sci. U.S.A.*, 103:14056–14061, 2006. 46

[86] B. S. Choudhuri, S. Bhakta, R. Barik, J. Basu, M. Kundu, and P. Chakrabati. Overexpression and functional characterization of an ABC (ATP-binding cassette) transporter encoded by the genes drrA and drrB of *Mycobacterium tuberculosis*. *Biochem. J.*, 367:279–285, 2002. 144

[87] S. Chung, S.-H. Shin, C. R. Bertozzi, and J. J. DeYoreo. Self-catalyzed growth of S layers via an amorphous-to-crystalline transition limited by folding kinetics. *Proc. Natl. Acad. Sci. U.S.A.*, 107:16,536–16,541, 2010. 81, 143

[88] ClassNK and N. K. Kyokai. Guidelines for container stowage and securing arrangements. [http://www.classnk.or.jp/hp/Publications/Publications_image/gl_container_e.pdf], 2009. 24

[89] M. I. Coates and M. J. Cohn. Vertebrate axial and appendicular patterning: the early development of paired appendages. *Amer. Zool.*, 39:676–685, 1999. 39

[90] M. J. Cohn and C. Tickle. Developmental basis of limblessness and axial patterning in snakes. *Nature*, 399:474–479, 1999. 40, 41

[91] C. Cole, J. D. Barber, and G. J. Barton. The Jpred 3 secondary structure prediction server. *Nucleic Acids Res.*, 36:W197–W201, 2008. 45

[92] W. Comper. *Extracellular Matrix.* Harwood Academic Publishers, 1996. 46

[93] C. E. Cook, M. L. Smith, M. J. Telford, A. Bastianello, and M. Akam. *Hox* genes and the phylogeny of the arthropods. *Curr. Biol.*, 11:759–763, 2001. 39

[94] J. Costerton. *The biofilm primer.* Springer, 2007. 33

[95] L. V. Crowley. *An introduction to human disease: pathology and pathophysiology Correlations.* Jones and Bartlett, 2009. 35

[96] R. A. Crowther, R. Henderson, and J. M. Smith. MRC image processing programs. *J. Struct. Biol.*, 116:9–16, 1996. 97

[97] M. E. Csete and J. C. Doyle. Reverse engineering biological complexity. *Science*, 295:1664–1669, 2002. 41, 42

[98] D. N. Wang and W. Kühlbrandt. Analysis of electron microscope images and electron diffraction patterns of thin crystals of Ø29 connectors in ice. *J. Mol. Biol.*, 217:691–699, 1991. 94, 109

[99] S. Dabholkar and M. Thattai. Brainstorming biology. *IET Synthetic Biology*, 1:17–20, 2007. 59

[100] S. A. Darst, H. O. Ribi, D. W. Pierce, and R. D. Kornberg. Two-dimensional crystals of *Escherichia coli* RNA polymerase holoenzyme on positively charged lipid layers. *J. Mol. Biol.*, 203:269–273, 1988. 91

[101] R. Das, P. J. Kiley, M. Segal, J. Norville, A. A. Y. AA, L. Wang, S. A. Trammell, L. E. Reddick, R. Kumar, F. Stellacci, N. Lebedev, J. Schnur, B. D. Bruce, S. Zhang, and M. Baldo. Integration of photosynthetic protein molecular complexes in solid-state electronic devices. *Nano. Letters*, 4:1079–1083, 2004. 112

[102] P. A. David and R. D. Stewart. *International logistics: the management of international trade operations.* Cengage Learning, 2010. 22

[103] D. G. Davies and C. N. H. Marques. A fatty acid messenger is responsible for inducing dispersion in microbial biofilms. *J. Bacteriol.*, 191:1393–1403, 2009. 34

[104] D. Densmore, T. H.-C. Hsiau, J. Kittleson, W. DeLoache, C. Batten, and J. C. Anderson. Algorithms for automated DNA assembly. *Nucleic Acids Res.*, 38: 2607–2616, 2010. 57

[105] R. Derda, A. Laromaine, A. Mammoto, S. K. Y. Tang, T. Mammoto, D. E. Ingber, and G. M. Whitesides. Paper-supported 3D cell culture for tissue-based bioassays. *Proc. Natl. Acad. Sci. U.S.A.*, 106:18457–18462, 2009. 117

[106] J. DiCarlo and J. Boeke. BBF RFC 38: Building blocks - standard large DNA genome construction. [http://dspace.mit.edu/handle/1721.1/49417], (Archived by WebCite at http://www.webcitation.org/5u3ZduEZT), *DSPACE*, 2009. 57

[107] R. Docampo, W. de Souza, K. Miranda, P. Rohloff, and S. N. J. Moreno. Acidocalcisomes–conserved from bacteria to man. *Nat. Rev. Micro.*, 3:251–261, 2005. 28

[108] R. F. Doolittle. The multiplicity of domains in proteins. *Annu. Rev. Biochem.*, 64:287–314, 1995. 46

[109] J. E. Dueber, G. C. Wu, G. R. Malmirchegini, T. S. Moon, C. J. Petzold, A. V. Ullal, K. L. J. Prather, and J. D. Keasling. Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.*, 27:753–761, 2009. 47, 89

[110] C. Dupraz and P. T. Visscher. Microbial lithification in marine stromatolites and hypersaline mats. *Trends Microbiol.*, 13:429–438, 2005. 31

[111] M. Dworkin and S. Falkow. *The Prokaryotes: Symbiotic associations, biotechnology, applied microbiology.* Springer, 2006. 117

[112] E. Egelseer, I. Schocher, M. Sara, and U. B. Sleytr. The S-Layer from *Bacillus stearothermophilus* DSM 2358 functions as an adhesion site for a high-molecular-weight amylase. *J. Bacteriol.*, 177:1444–1451, 1995. 29

[113] T. Ellis, X. Wang, and J. J. Collins. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat. Biotechnol.*, 27:465–471, 2009. 58

[114] M. Ellison, D. Ridgway, J. Fedor, E. Garside, and K. R. D. Lloyd. BBF RFC 47: BioBytes assembly standard. [`http://dspace.mit.edu/handle/1721.1/49518`], (Archived by WebCite at `http://www.webcitation.org/5u3ZTTONi`), *DSPACE*, 2009. 57

[115] M. B. Elowitz, M. G. Surette, P.-E. Wolf, J. B. Stock, and S. Leibler. Protein mobility in cytoplam of *Escherichia coli*. *J. Bacteriol.*, 181:197–203, 1999. 29

[116] L. M. Elson, C. Vadala, and J. Giuffré. *The zoology coloring book.* Harper-Perennial, 1982. 38

[117] J. D. Enderle and J. D. Bronzino. *Introduction to biomedical engineering.* Academic Press, 2011. 27

[118] D. Endy. Foundations for engineering biology. *Nature*, 438:449–453, 2005. 57

[119] H. Engelhardt and J. Peters. Structural research on S-layers: a focus on stability, surface layer homology domains, and surface layer-cell wall interactions. *J. Struct. Biol.*, 124:276–302, 1998. 81

[120] H. Engelhardt, W. O. Saxton, and W. Baumeister. Three-dimensional structure of the tetragonal surface layer of *Sporosarcina ureae*. *J. Bacteriol.*, 168:309–317, 1986. 112

[121] C. Engler, R. Gruetzner, R. Kandzia, and S. Marillonnet. Golden gate shuffling: a one-pot DNA shuffling method based on Type IIs restriction enzymes. *PLoS ONE*, 4:e5553, 2009. 57

[122] J. Errington. Dynamic proteins and a cytoskeleton in bacteria. *Nat. Cell Biol.*, 5:175–175, 2003. 27

[123] M. A. Fedonkin. *The rise of animals: evolution and diversification of the kingdom animalia.* John Hopkins University Press, 2007. 38

[124] S. Ferro-Novick and R. Jahn. Vesicle fusion from yeast to man. *Nature*, 370: 191–193, 1994. 35

[125] J. R. Finnerty, K. Pang, P. Burton, D. Paulson, and M. Q. Martindale. Origins of bilateral symmetry: *Hox* and *Dpp* expression in a sea anemone. *Science*, 304: 1335–1337, 2004. 39

[126] H.-C. Flemming, T. R. Neu, and D. J. Wozniak. The EPS matrix: the "house of biofilm cells". *J. Bacteriol*, 189:7945–7947, 2007. 31

[127] E. P. Foley. *The Law of Life and Death.* Harvard University Press, 2011. 27

[128] P. Fromherz. Electron microscopic studies of lipid protein films. *Nature*, 231: 267–268, 1971. 91

[129] M. Galfione, W. Luo, J. Kim, D. Hawke, R. Kobayashi, C. Clapp, L.-Y. Yu-Lee, and S.-H. Lin. Expression and purification of the angiogenesis inhibitor 16-kDa prolactin fragment from insect cells. *Protein Expression Purif.*, 28:252–258, 2003. 82, 85, 144

[130] E. D. Gardner, D. J. Gray, and R. O'Rahilly. *Anatomy; a regional study of human structure.* Saunders, 1969. 40

[131] T. S. Gardner, C. R. Cantor, and J. J. Collins. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 403:339–342, 2000. 75

[132] S. J. Gaunt. Expression patterns of mouse Hox genes: clues to an understanding of developmental and evolutionary strategies. *BioEssays*, 13:505–513, 1991. 40

[133] G. G. Geesey. Bacterial behavior at surfaces. *Curr. Opin. Microbiol.*, 4:296–300, 2001. 33

[134] W. J. Gehring, M. Affolter, and T. Bürglin. Homeodomain proteins. *Annu. Rev. Biochem.*, 63:487–526, 1994. 38

[135] G. Georgiou and P. Valax. Expression of correctly folded proteins in *Escherichia coli*. *Curr. Opin. Biotech.*, 7:190–197, 1996. 144

[136] G. D. Geske, J. C. O'Neill, and H. E. Blackwell. Expanding dialogues: from natural autoinducers to non-natural analogues that modulate quorum sensing in Gram-negative bacteria. *Chem. Soc. Rev.*, 37:1432–1447, 2008. 116

[137] D. Gibson, G. A. Benders, C. Andrews-Pfannkoch, E. A. Denisova, H. Baden-Tillson, J. Zaveri, T. B. Stockwell, A. Brownley, D. W. Thomas, M. A. Algire, C. Merryman, L. Young, V. N. Noskov, J. I. Glass, J. C. Venter, C. A. H. III, and H. O. Smith. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science*, 319:1215–1220, 2008. 45, 57

[138] D. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. H. III, and H. O. Smith. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, 6:343–345, 2009. 57

[139] N. W. Gillham, J. E. Boynton, and G. R. Hauser. Translational regulation of gene expression in chloroplasts and mitochondria. *Annu. Rev. Genet.*, 28:71–93, 1994. 36

[140] B. Gipson, X. Z. Z. Y. Zhang, and H. Stahlberg. 2dx–user friendly image processing for 2D crystals. *J. Struct. Biol.*, 157:64–72, 2007. 85, 97

[141] C. Gomez, E. M. Ozbudak, J. Wunderlich, D. Baumann, J. Lewis, and O. Pourquie. Control of segment number in vertebrate embryos. *Nature*, 454: 335–339, 2008. 40

[142] T. Gonen, Y. Cheng, P. Sliz, Y. Hiroaki, Y. Fujiyoshi, S. C. Harrison, and T. Walz. Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature*, 438:633–638, 2005. 93

[143] S. Goodman. *Medical cell biology.* Elsevier/Academic Press, 2008. 37

[144] G. Griffiths. On vesicles and membrane compartments. *Protoplasma*, 195:37–58, 1996. 35

[145] R. Grunberg. BBF RFC 24: Conversion of Freiburg (Fusion) Biobricks to the Silver (BioFusion) format. [`http://dspace.mit.edu/handle/1721.1/44961`], (Archived by WebCite at `http://www.webcitation.org/5u3YcCmXe`), *DSPACE*, 2009. 47, 57

[146] B. K. S. Gupta. *Modern foraminifera.* Kluwer Academic Publishers, 2002. 37

[147] N. Gyobu, K. Tani, Y. Hiroaki, A. Kamegawa, K. Mitsuoka, and Y. Fujiyoshi. Improved specimen preparation for cryo-electron microscopy using a symmetric carbon sandwich technique. *J. Struct. Biol.*, 146:325–333, 2004. 94, 104

[148] H.-B. Zhang and L.-H. Wang and L.-H. Zhang. *Detection and analysis of quorum-quenching enzymes against acyl homoserine lactone quorum-sensing signals.* John Wiley Sons, Inc., 2005. 116

[149] S. E. Halford, A. J. Welsh, and M. D. Szczelkun. Enzyme-mediated DNA looping. *Annu. Rev. Biophys. BIomol. Struct.*, 33:1–24, 2004. 60

[150] L. Hall-Stoodley, J. W. Costerton, and P. Stoodley. Bacterial biofilms: from the natural environment to infectious diseases. *Microbiol.*, 107:95–108, 2004. 116

[151] E. L. Haseltine and F. H. Arnold. Synthetic gene circuits: design with directed evolution. *Annu. Rev. Biophys. Biomol. Struct.*, 36:1–19, 2007. 58

[152] K. A. Haynes, M. L. Broderick, A. D. Brown, T. L. Butner, J. O. Dickson, W. L. Harden, L. H. Heard, E. L. Jessen, K. J. Malloy, B. J. Ogden, S. Rosemond,

S. Simpson, E. Zwack, A. M. Campbell, T. T. Eckdahl, L. J. Heyer, and J. L. Poet. Engineering bacteria to solve the Burnt Pancake Problem. *J. Biol. Eng.*, 2:8, 2008. 59

[153] R. Henderson and P. N. T. Unwin. Three-dimensional model of purple membrane obtained by electron microscopy. *Nature*, 257:28–32, 1975. 109

[154] R. Henderson, J. M. Baldwin, K. H. Downing, J. Lepault, and F. Zemlin. Structure of purple membrane from *Halobacterium halobium*: recording, measurement and evaluation of electron micrographs at 3.5 Å resolution. *Ultramicroscopy*, 19:147–178, 1986. 106

[155] R. Henderson, J. M. Baldwin, T. A. Ceska, F. Zemlin, E. Beckmann, and K. H. Downing. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.*, 213:899–929, 1990. 93, 109

[156] E. C. Hett and E. J. Rubin. Bacterial growth and cell division: a mycobacterial perspective. *Microbiol. Mol. Biol. Rev.*, 72:126–156, 2008. 28

[157] A. Hill. *The encyclopedia of operations management; A field manual and glossary of operations management terms and concepts.* Pearson Education, 2011. 23, 24

[158] T. Hirai, K. Murata, K. Mitsuoka, Y. Kimura, and Y. Fujiyoshi. Trehalose embedding technique for high-resolution electron crystallography: application to structural study on bacteriorhoposin. *J. Electron Microsc.*, 48:653–658, 1999. 94, 104, 109

[159] Y. Hiroaki, K. Tani, A. Kamegawa, N. Gyobu, K. Nishikawa, Suzuki, T. Walz, S. Sasaki, K. Mitsuoka, K. Kimura, A. Mizoguchi, and Y. Fujiyoshi. Implications of the aquaporin-4 structure on array formation and cell adhesion. *J. Mol. Biol.*, 355:628–639, 2006. 93

[160] P. J. Holm, P. Bhakat, C. Jegerschold, N. Gyobu, K. Mitsuoka, Y. Fujiyoshi,

R. Morgenstern, and H. Hebert. Structural basis for detoxification and oxidative stress protection in membranes. *J. Mol. Biol.*, 360:934–945, 2006. 93

[161] W. J. Holtz and J. D. Keasling. Engineering static and dynamic control of synthetic pathways. *Cell*, 140:19–23, 2010. 57

[162] J. Hou, G. E. Sims, C. Zhang, and S.-H. Kim. A global representation of the protein fold space. *Proc. Natl. Acad. Sci. U.S.A.*, 100:2386–2390, 2003. 46

[163] H.-H. Huang, D. Camsund, P. Lindblad, and T. Hieidorn. Design and characterization of molecular tools for a Synthetic Biology approach to developing cyanobacterial biotechnology. *Nucleic Acids Res.*, 38:2727–2735, 2010. 59

[164] G. W. Huisman and R. Kolter. Sensing starvation: a homoserine lactone–dependent signaling pathway in *Escherichia coli*. *Science*, 265:537–539, 1994. 134

[165] G. V. Hulme. RFID helps to track cargo containers. *InformationWeek*, April 6, 2005. 26

[166] N. Ilk, P. Kosma, M. Puchberger, E. M. Egelseer, H. F. Mayer, U. B. Sleytr, and M. Sara. Structural and functional analyses of the secondary cell wall polymer of *Bacillus sphaericus* CCM 2177 that serves as an S-layer-specific anchor. *J. Bacteriol.*, 181:7643–7646, 1999. 81, 143

[167] P. Ioannou. *Intelligent freight transportation.* Automation and control engineering. CRC Press, 2008. 22

[168] M. Itaya, K. Fujita, A. Kuroki, and K. Tsuge. Bottom-up genome assembly using the *Bacillus subtilis* genome vector. *Nat. Methods*, 6:343–345, 2009. 57

[169] J. Tang and A. Ebner and H. Badelt-Lichtblau and C. Völlenkle and C. Rankl and B. Kraxberger and M. Leitner and L. Wildling and H. J. Gruber and U. B. Sleytr and N. Ilk and P. Hinterdorfer. Recognition imaging and highly ordered

molecular templating of bacterial S-layer nanoarrays containing affinity-tags. *Nano. Lett.*, 8:4312–4319, 2008. 82

[170] W. Jakob, S. Sagasser, S. Dellaporta, P. Holland, K. Kuhn, and B. Schierwater. The *Trox*-2 Hox/ParaHox gene of *Trichoplax* (Placozoa) marks an epithelial boundary. *Dev. Genes Evol.*, 214:170–175, 2004. 39

[171] J. Jansson and D. Shneerson. *Port economics.* Transportation Studies. MIT Press, 1982. 25

[172] B. K. Jap and K. H. Downing. Structure of PhoE porin in projection at 3.5 Å resolution. *J. Struct. Biol.*, 103:57–63, 1990. 94, 104, 109

[173] M. Jarosch, E. M. Egelseer, C. Huber, D. Moll, D. Mattanovich, U. B. Sleytr, and M. Sara. Analysis of the structure-function relationship of the S-layer protein SbsC of *Bacillus stearothermophilus* ATCC 12980 by producing truncated forms. *Microbiol.*, 147:1353–1363, 2001. 144

[174] A. Jayaraman, A. K. Sun, and T. Wood. Characterization of axenic *Pseudomonas fragi* and *Escherichia coli* biofilms that inhibit corrosion of SAE 1018 steel. *J. Appl. Microbiol.*, 84:485–492, 1998. 117

[175] H. Joo, Z. Lin, and F. A. Arnold. Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation. *Nature*, 399:670–673, 1999. 60

[176] D. E. K. Jr. The seven pillars of life. *Science*, 295:2215–2216, 2002. 27

[177] R. J. P. Jr. and D. C. White. Developmental biology of biofilms implications for treatment and control. *Trends Microbiol.*, 5:435–440, 1997. 116

[178] T. K. Jr. Fundamental principles for engineering biological systems. [`http://csbi.mit.edu/people/knight.html`], April 16, 2008. 142

[179] H.-Q. Juang, N. A. Bos, and J. J. Cebra. Timing, localization, and persistence of colonization by segmented filamentous bacteria in the neonatal mouse gut

depend on immune status of mothers and pups. *Infect. Immun.*, 69:3611–3617, 2001. 32

[180] S. Jurenaite-Urbanaviciene, J. Serksnaite, E. Kriukiene, J. Giedriene, C. Venclovas, and A. Lubys. Generation of DNA cleavage specificities of Type II restriction endonucleases by reassortment of target recognition domains. *Proc. Natl. Acad. Sci. U.S.A.*, 104:10358–10363, 2007. 60

[181] A. K. Kamrani and E. A. Nasr. *Engineering design and rapid prototyping.* Spinger, New York, NY, 2010. 21

[182] O. Kandror, A. DeLeon, and A. L. Goldberg. Trehalose synthesis is induced upon exposure of *Escherichia coli* to cold and is essential for viability at low temperatures. *Proc. Natl. Acad. Sci. U.S.A.*, 99:9727–9732, 2002. 109

[183] A. Kaplan and L. Reinhold. CO_2 concentrating mechanisms in photosynthetic microorganisms. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 50:539–570, 1999. 28

[184] G. Karleskint, R. Turner, and J. Small. *Introduction to Marine Biology.* Brooks/Cole, 2009. 31

[185] G. Karp. *Cell and molecular biology: concepts and experiments.* John Wiley & Sons, 2009. 37

[186] C. N. Keim, J. L. Martins, F. Abreu, A. S. Rosado, H. L. de Barros, R. Borojevic, U. Lins, and M. Farina. Multicellular life cycle of magnetotactic prokaryotes. *FEMS Microbiol. Lett.*, 240:203–208, 2004. 31, 32, 33, 34

[187] D. F. Kelly and K. A. Taylor. Identification of the β1-integrin binding site on α-actinin by cryoelectron microscopy. *J. Struct. Biol.*, 149:290–302, 2005. 92

[188] D. F. Kelly, D. W. Taylor, C. Bakolitsa, A. A. Bobkov, L. Bankston, R. C. Liddington, and K. A. Taylor. Structure of the α-actinin-vinculin head domain

complex determined by cryo-electron microscopy. *J. Mol. Biol.*, 357:562–573, 2006. 92, 93

[189] D. F. Kelly, D. Dukovski, and T. Walz. Monolayer purification: a rapid method for isolating protein complexes for single-particle electron microscopy. *Proc. Natl. Acad. Sci. U.S.A.*, 105:4703–4708, 2008. 76, 83

[190] J. R. Kelly, A. J. Rubin, J. H. Davis, C. M. Ajo-Franklin, J. Cumbers, M. J. Czar, K. de Mora, A. L. Glieberman, D. D. Monie, and D. Endy. Measuring the activity of BioBrick promoters using an in vivo reference standard. *J. Biol. Eng.*, 3:4, 2009. 59

[191] C. Kepplinger, N. Ilk, U. B. Sleytr, and B. Schuster. Intact lipid vesicles reversibly tethered to a bacterial S-layer protein lattice. *Soft Matter*, 5:325–333, 2008. 82, 143

[192] Y.-G. Kim, J. Cha, and S. Chandrasegaran. Hybrid restriction enzymes: zinc finger fusions to FokI cleavage domain. *Proc. Natl. Acad. Sci. U.S.A.*, 93:1156–1160, 1996. 60

[193] Y. Kimura, D. G. Vassylyev, A. Miyazawa, A. Kidera, M. Matsushima, K. Mitsuoka, K. Murata, T. Hirai, and Y. Fujiyoshi. Surface of bacteriorhodopsin revealed by high-resolution electron crystallography. *Nature*, 389:206–211, 1997. 93

[194] T. F. Knight. BBF RFC 10: Draft standard for BioBrick biological parts. [http://dspace.mit.edu/handle/1721.1/45138], (Archived by WebCite at http://www.webcitation.org/5u3Xs6OL6), *DSPACE*, 2007. 45, 47, 57

[195] T. F. Knight. BBF RFC 12: Draft BioBrick BB-2 standard for biological parts. [http://dspace.mit.edu/handle/1721.1/45139], (Archived by WebCite at http://www.webcitation.org/5u3XyaQ9i), *DSPACE*, 2008. 47, 57

[196] A. L. Koch. Were Gram-positive rods the first bacteria? *Trends Microbiol.*, 11: 166–169, 2003. 33

[197] A. Komeili, Z. Li, D. K. Newman, and G. J. Jensen. Magentosomes are cell membrane invaginations organized by the actin-like protein MamK. *Science*, 311:242–245, 2006. 28, 30

[198] D. S. Kong, P. A. Carr, L. Chen, S. Zhang, and J. M. Jacobson. Parallel gene synthesis in a microfluidic device. *Nucl. Acids Res.*, 35:e61, 2007. 45

[199] W. N. Konings, K. J. Hellingwerf, and G. T. Robillard. *Membrane transport*, chapter Transport across bacterial membranes, pages 257–284. New comprehensive biochemistry. Elsevier/North-Holland Biomedical Press, 1981. 27, 29

[200] E. V. Koonin, Y. I. Wolf, and G. P. Karev. The structure of the protein universe and genome evolution. *Nature*, 420:218–223, 2002. 46

[201] S. Krawetz. *Bioinformatics for systems biology.* Springer, 2009. ISBN 9781934115022. URL `http://books.google.com/books?id=s0SxfCuJfOsC`. 37

[202] W. E. Krumbein, D. M. Paterson, and G. G. A. Zavarzin. *Fossil and recent biofilms: a natural history of life on Earth.* Kluwer Academic Publishers, 2003. 32, 34

[203] E. W. Kubalek, R. D. Kornberg, and S. A. Darst. Improved transfer of two-dimensional crystals from the air/water interface to specimen support grids for high-resolution analysis by electron microscopy. *Ultramicroscopy*, 35:295–304, 1991. 92, 93, 101, 109, 110, 111

[204] E. W. Kubalek, S. F. L. Grice, and P. O. Brown. Two-dimensional crystallization of histidine-tagged, HIV-1 reverse transcriptase promoted by a novel nickel-chelating lipid. *J. Struct. Biol.*, 113:117–123, 1994. 91

[205] L. Lebeau and F. Lach and C. Venien-Bryan and A. Renault and J. Dietrich and T. Jahn and M. G. Palmgren and W. Kühlbrandt and C. Mioskowski. Two-dimensional crystallization of a membrane protein on a detergent-resistant lipid monolayer. *J. Mol. Biol.*, 308:639–647, 2001. 92

[206] R. Lamed, J. Naimark, E. Morgenstern, and E. A. Bayer. Specialized cell surface structures in cellulolytic bacteria. *J. Bacteriol.*, 169:3792 – 3800, 1987. 30

[207] I. Langmuir and V. J. Schaefer. Activities of urease and pepsin monolayers. *J. Am. Chem. Soc.*, 60:1351–1360, 1938. 93

[208] M. Leabu. Membrane fusion in cells: molecular machinery and mechanisms. *J. Cell. Mol. Med.*, 10:423–427, 2006. 36

[209] H. Ledford. Giant bacterium carries thousands of genomes. *Nature*, page doi:10.1038/news.2008.806, 2008. 27

[210] C. T. Lefevre, F. Abreu, U. Lins, and D. A. Bazylinski. Nonmagnetotactic multicellular prokaryotes from low-saline, nonmarine aquatic environments and their unusual negative phototactic behavior. *Appl. Environ. Microbiol.*, 76: 3220–3227, 2010. 32, 34

[211] E. Leonard, P. K. Ajikumar, K. Thayer, W.-H. Xiao, J. D. Mo, B. Tidor, G. Stephanopoulos, and K. L. J. Prather. Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. *Proc. Natl. Acad. Sci. U.S.A.*, 107:13654–13659, 2010. 47

[212] J. Lepault and T. Pitt. Projected structure of unstained, frozen-hydrated T-layer of *Bacillus brevis*. *EMBO J.*, 3:101–105, 1984. 94, 105, 108, 111

[213] J. Lepault, N. Martin, and K. Leonard. Three-dimensional structure of the T-layer of *Bacillus sphaericus* P-1. *J. Bacteriol.*, 168:303–308, 1986. 94, 105, 108, 111

[214] M. Levinson. *The box: how the shipping container made the world smaller and the world economy bigger.* Princeton University Press, Princeton, N.J., 2006. 23, 25

[215] A. Levskaya, A. A. Chevalier, J. J. Tabor, Z. B. Simpson, L. A. Lavery, M. Levy, E. A. Davidson, A. Scouras, A. D. Ellington, E. M. Marcotte, and C. A. Voigt. Synthetic biology: engineering *Escherichia coli* to see light. *Nature*, 438:441–442, 2005. 59

[216] D. Levy, G. Mosser, O. Lambert, G. S. Moeck, D. Bald, and J. L. Rigaud. Two-dimensional crystallization on lipid layer: a successful approach for membrane proteins. *J. Struct. Biol.*, 127:44–52, 1999. 92

[217] B. Lewin. *Cells.* Jones and Bartlett Publishers, 2007. 37

[218] P. J. Lewis, S. D. Thaker, and J. Errington. Compartmentalization of transcription and translation in *Bacillus subtilis*. *EMBO J.*, 19:710–718, 2000. 29

[219] H. Li, C. Tang, and N. S. Wingreen. Are protein folds atypical? *Proc. Natl. Acad. Sci. U.S.A.*, 95:4987–4990, 1998. 46

[220] M. Li and S. J. Elledge. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat. Methods*, 4:251–256, 2007. 57

[221] Z. Li and G. J. Jensen. Electron cryotomography: a new view of ultrastructure. *Curr. Opin. Microbiol.*, 12:333–340, 2009. 28

[222] Z. Li, R. K. Hite, Y. Cheng, and T. Walz. Evaluation of imaging plates as recording medium for images of negatively stained single particles and electron diffraction patterns of two-dimensional crystals. *J. Electron Microsc.*, 59:53–63, 2010. 85

[223] J. Liu, T. Wendt, D. Taylor, and K. Taylor. Refined model of the 10 S conformation of smooth muscle myosin by cryo-electron microscopy 3D image reconstruction. *J. Mol. Biol.*, 329:963–972, 2003. 93

[224] J. Liu, D. W. Taylor, and K. A. Taylor. A 3-D reconstruction of smooth muscle α-actinin by cryo-EM reveals two different conformations at the actin-binding region. *J. Mol. Biol.*, 338:115–125, 2004. 93

[225] J. Liu, D. W. Taylor, E. B. Krementsova, K. M. Trybus, and K. A. Taylor. Three-dimensional structure of the myosin V inhibited state by cryo-electron tomography. *Nature*, 442:208–211, 2006. 93

[226] P. M. Llopis, A. F. Jackson, O. Sliusarenko, I. Surovtsev, J. Heinritz, T. Emonet, and C. Jacobs-Wagner. Spatial organization of the flow of genetic information in bacteria. *Nature*, 466:77–81, 2010. 29

[227] H. F. Lodish. *Molecular cell biology.* W.H. Freeman, 2008. 35

[228] J. W. Long, B. Dunn, D. R. Rolison, and H. S. White. Three-dimensional battery architectures. *Chem. Rev.*, 104:4463–4492, 2004. 112

[229] D. Lopez, H. Vlamakis, and R. Kolter. Biofilms. *Cold Spring Harb. Perspect. Biol.*, 2:a000398, 2010. 32

[230] C. J. Lowe, M. Terasaki, M. Wu, R. M. F. Jr., L. Runft, K. Kwan, S. Haigo, J. Aronowicz, E. Lander, G. Gruber, M. Smith, M. Kirschner, and J. Gerhart. Dorsoventral patterning in hemichordates: insights into early chordate evolution. *PLoS*, 4:e291, 2006. 40

[231] T. K. Lu and J. J. Collins. Dispersing biofilms with engineered enzymatic bacteriophage. *Proc. Natl. Acad. Sci. U.S.A.*, 104:11197–11202, 2007. 117

[232] T. K. Lu and J. J. Collins. Morphogen-defined patterning of *Escherichia coli* enabled by an externally tunable band-pass filter. *J. Biol. Eng.*, 3:10, 2009. 117

[233] T. K. Lu, A. S. Khalil, and J. J. Collins. Next generation synthetic gene networks. *Nat. Biotechnol.*, 27:946–950, 2009. 58

[234] M. S. Luijsterburg, M. C. Noom, G. J. L. Wuite, and R. T. Dame. The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J. Struct. Biol.*, 156:262–272, 2006. 29

[235] S. R. Lybarger and J. R. Maddock. Polarity in action: asymmetric protein localization in bacteria. *J. Bacteriol.*, 183:3261–3267, 2001. 27

[236] L. MacKenzie, D. K. Arwine, E. J. Shewan, and M. J. McHugh. *Biology: A Search For Order In Complexity.* Christian Liberty Press, 2005. 38

[237] T. Maniatis, E. F. Fritsch, and J. Sambrook. *Molecular cloning: a laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1982. 57

[238] J. Marshall, D. M. Gowers, and S. E. Halford. Restriction endonucleases that bridge and excise two recognition sites from DNA. *J. Mol. Biol.*, 367:419–431, 2007. 60

[239] L. Martin, A. Che, and D. Endy. Gemini, a bifunctional enzymatic and fluorescent reporter of gene expression. *PLoS ONE*, 4:e7569, 2009. 59

[240] W. Martin and R. G. Herrmann. Gene transfer from organelles to the nucleus: how much, what happens, and why? *Plant Physiol.*, 118:9–17, 1998. 36

[241] B. Martoglio and B. Dobberstein. Signal sequences: more than just greasy peptides. *Trends in Cell Biology*, 8:410 – 415, 1998. 30

[242] A. J. Mayo and N. Nohria. *In their time : the greatest business leaders of the twentieth century.* Harvard Business School Press, Boston MA, 2005. 22

[243] R. McDaniel and R. Weiss. Advances in synthetic biology: on the path from prototypes to applications. *Curr. Opin. Biotech.*, 16:476–483, 2005. 57

[244] M. McNicholas. *Maritime security: an introduction.* Butterworth–Heinemann homeland security series. Butterworth–Heinemann, 2007. 23, 25, 26

[245] J. E. Mendell, K. D. Clements, J. H. Choat, and E. R. Angert. Extreme polyploidy in a large bacterium. *Proc. Natl. Acad. Sci. U.S.A.*, 105:6703–6734, 2008. 27

[246] A. Meyer and E. Malaga-Trillo. Vertebrate genomics: more fishy tales about *Hox* genes. *Curr. Biol.*, 9:210–213, 1999. 39

[247] B. Moore. RFID: cargo security and product traceability. *Association for Automatic Identification and Mobility*, March 3, 2010. 26

[248] D. M. Morris and G. J. Jensen. Toward a biomechanical understanding of whole bacterial cells. *Annu. Rev. Biochem.*, 77:583–613, 2008. 28

[249] G. Mosser, C. Ravanat, J. M. Freyssinet, and A. Brisson. Sub-domain structure of lipid-bound annexin-V resolved by electron image analysis. *J. Mol. Biol.*, 217: 241–245, 1991. 91

[250] E. Mowatt-Larssen. *TankTainer: a portable bulk liquid tank container for intermodal service.* 1978. 24

[251] A. Mukhopadhyay, A. M. Redding, B. J. Rutherford, and J. D. Keasling. Importance of systems biology in engineering microbes for biofuel production. *Curr. Opin. Biotech.*, 19:228–234, 2008. 57, 89

[252] K. Murata, K. Mitsuoka, T. Hirai, T. Walz, P. Agre, J. B. Heymann, A. Engel, and Y. Fujiyoshi. Structural determinants of water permeation through aquaporin-1. *Nature*, 407:599–605, 2000. 93, 109

[253] C. Murphy and J. Yates. *The International Organization for Standardization (ISO): global governance through voluntary consensus.* Global institutions series. Routledge, 2009. 23

[254] J. Murray. *Ecology and applications of benthic foraminifera.* Cambridge University Press, 2006. 36

[255] R. G. E. Murray, M. Hall, and B. G. Thompson. Cell division in *Deinococcus radiodurans* and a method for displaying septa. *Can. J. Microbiol.*, 29:1412–1423, 1983. 32

[256] K. Muller, K. M. Arndt, iGEM 2007, and R. Grunberg. BBF RFC 25: Fusion protein (Freiburg) Biobrick assembly standard. [http://dspace.mit.edu/handle/1721.1/45140], (Archived by WebCite at http://www.webcitation.org/5u3YhFG3y), *DSPACE*, 2009. 47, 57

[257] N. Ilk and C. Völlenkle and E. M. Egelseer and A. Breitwieser and U. B. Sleytr and M. Sara. Molecular characterization of the S-layer gene, SbpA, of *Bacillus sphaericus* CCM 2177 and production of a functional S-layer fusion protein with the ability to recrystallize in a defined orientation while presenting the fused allergen. *Appl. Environ. Microbiol.*, 68:3251–3260, 2002. 16, 81, 82, 85, 87, 143, 144, 145, 146

[258] K. T. Nam, D.-W. Kim, P. J. Yoo, C.-Y. Chiang, N. Meethong, P. T. Hammond, Y.-M. Chiang, and A. M. Belcher. Virus-enabled synthesis and assembly of nanowires for lithium ion battery electrodes. *Science*, 312:885–888, 2006. 112

[259] A. Navarrete, A. Peacock, S. J. Macnaughton, J. Urmeneta, J. Mas-Castella, D. C. White, and R. Guerrero. Physiological status and community composition of microbial mats of the Ebro Delta, Spain, by signature lipid biomarkers. *Microb. Ecol.*, 39:92–99, 2000. 33, 34

[260] H. Niki, Y. Yamaichi, and S. Hiraga. Dynamic organization of chromosomal DNA in *Escherichia coli*. *Genes Dev.*, 14:212–223, 2000. 29

[261] S. Normark, H. G. Boman, and E. Matsson. Mutant of *Escherichia coli* with anomalous cell division and ability to decrease episomally and chromosomally mediated resistance to ampicillan and several other antibiotics. *J. Bacteriol.*, 97:1334–1342, 1969. 32

[262] J. Norville, A. Belcher, and T. F. K. Jr. BBF RFC 15: Innovations mean nothing unless you use them the new BioScaffold family of BioBrick parts to enable manipulations such as protein fusions, library construction, and part domestication. [http://bbf.openwetware.org/BBFRFC15.html], (Archived by WebCite at http://www.webcitation.org/5u3YPWzw3), *DSPACE*, 2008. 57

[263] J. E. Norville, D. F. Kelly, T. F. K. Jr., A. M. Belcher, and T. Walz. 7 Å projection map of the S-layer protein sbpA obtained with trehalose-embedded monolayer crystals. *J. Struct. Biol.*, 160:313–323, 2007. 6, 80, 81, 82, 87, 143, 144, 145, 146, 150, 151

[264] J. E. Norville, R. Derda, S. Gupta, K. A. Drinkwater, A. M. Belcher, and T. F. K. Jr. Introduction of customized inserts for streamlined assembly and optimization of BioBrick synthetic genetic circuits. *J. Biol. Eng.*, 4:17, 2010. 5, 56, 89, 133, 149

[265] J. E. Norville, D. F. Kelly, T. F. K. Jr., A. M. Belcher, and T. Walz. Fast and easy protocol for the purification of recombinant S-layer protein for synthetic biology applications. *Biotechnol. J.*, 6:807–811, 2011. 6, 80, 143

[266] J. E. Norville, S. Pennybaker, R. Derda, M. R. Lockett, C. Li, G. Whitesides, A. M. Belcher, and T. F. K. Jr. Paper supported polycultures of synthetically programmed bacteria. *in preparation*, 2012. 6

[267] M. Ohi, Y. Li, Y. Cheng, and T. Walz. Negative staining and image classification – powerful tools in modern electron microscopy. *Biol. Proced. Online*, 6: 23–34, 2004. 83

[268] K. W. Osteryoung and J. Nunnari. The division of endosymbiotic organelles. *Science*, 302:1698–1704, 2003. 36

[269] H. W. Paerl, J. L. Pinckney, and T. F. Steppe. Cyanobacterial-bacterial mat

consortia: examining the functional unit of microbial survival and growth in extreme environments. *Environ. Microbiol.*, 2:11–26, 2000. 31

[270] D. Papapostolou and S. Howorka. Engineering and exploiting protein assemblies in synthetic biology. *Mol. BioSyst.*, 5:723–732, 2009. 81

[271] J. B. Parsons, S. Frank, D. Bhella, M. Liang, M. B. Prentice, D. P. Mulvihill, and M. J. Warren. Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement. *Mol. Cell*, 38:305–315, 2010. 30

[272] D. M. Paterson, R. J. Aspden, and R. P. Reid. *Microbial mats: modern and ancient microorganisms in stratified systems*, chapter Biodynamics of modern marine stromatolites, pages 225–238. Cellular origin, life in extreme habitats and astrobiology. Springer, 2010. 27

[273] J. C. Pearson, D. Lemons, and W. McGinnis. Modulating Hox gene functions during animal body patterning. *Nat. Rev. Genet.*, 6:893–904, 2005. 38, 39, 142

[274] S. Peisajovich, A. Horwitz, O. Holler, B. Rhau, and W. Lim. BBF RFC 28: A method for combinatorial multi-part assembly based on the Type IIs restriction enzyme AarI. [http://dspace.mit.edu/handle/1721.1/46721], (Archived by WebCite at http://www.webcitation.org/5u3ZYki0A), *DSPACE*, 2009. 57

[275] A. C. Perez-Osorio and M. J. Franklin. Isolation of RNA and DNA from biofilm samples obtained by laser capture microdissection microscopy. *Cold Spring Harb. Protoc.*, 2008. doi: http://dx.doi.org/10.1101/pdb.prot5065. 116

[276] I. E. Phillips and P. A. Silver. BBF RFC 23: A new BioBrick assembly strategy designed for facile protein engineering. [http://dspace.mit.edu/handle/1721.1/32535], (Archived by WebCite at http://www.webcitation.org/5u3YC39L5), *DSPACE*, 2006. 47, 57, 59

[277] R. B. Phillips, J.Kondev, and J. Theriot. *Physical biology of the cell.* Garland Science, 2009. 117

[278] F. Pichaud and C. Desplan. Pax genes and eye organogenesis. *Curr. Opin. Genet. Dev.*, 12:430–434, 2002. 40

[279] A. Pollack. Custom-made microbes, at your service. *New York Times*, January 17, 2006. 44

[280] J. Pommerville. *Alcamo's Fundamentals of Microbiology.* Jones & Bartlett Learning, 2010. 29

[281] C. P. Ponting and R. R. Russell. The natural history of protein domains. *Annu. Rev. Biophys. Biomol. Struct.*, 31:45–71, 2002. 46, 47

[282] O. Pourquie. The segmentation clock: converting embryonic time into spatial pattern. *Science*, 301:328–330, 2003. 40

[283] K. L. J. Prather and A. Jaramillo. Focus on synthetic biology. *Biotechnol. J.*, 4:1367, 2009. 81

[284] C. C. Project. The CCP4 suite: programs for protein crystallography. *Acta. Crystallogr. D Biol. Crystallogr.*, 50:760–763, 1994. 97

[285] C. Provost. *Grow Your Profits.* iUniverse Incorporated, 2010. 22

[286] D. Pum and U. B. Sleytr. Large-scale reconstitution of crystalline bacterial surface-layer proteins at the air-water-interface and on lipid films. *Thin Solid Films*, 244:882–886, 1994. 94, 99, 111

[287] D. Pum and U. B. Sleytr. Anisotropic crystal growth of the S-layer of *Bacillus sphaericus* CCM 2177 at the air/water interface. *Colloid. Surface. A*, 102: 99–104, 1995. 94, 97

[288] D. Pum, M. Sara, and U. B. Sleytr. Structure, surface charge, and self-assembly of the S-layer lattice from *Bacillus coagulans* E3866. *J. Bacteriol.*, 171:5296–5303, 1989. 87

[289] G. Qing, L.-C. Ma, A. Khorchid, G. V. T. Swapna, T. K. Ma, M. M. Takayama, B. Xia, S. Phadtare, H. Ke, T. Acton, G. T. Montelione, M. Ikura, and M. Inouye. Cold-shock induced high-yield protein production in *Escherichia coli*. *Nat. Biotechnol.*, 22:877–882, 2004. 144

[290] R. Quek, K. Lim, and Y. Chui. *Contemporary ergonomics*, chapter Visual displays: the effect of two- and three-dinensional displays on remote crane control performance, pages 419–428. Contemporary Ergonomics. Taylor & Francis, 2000. 25

[291] S. M. Rapoport, T. Schewe, and B.-J. Thiele. *Erythroid cells*, chapter Maturational breakdown of mitochondria and other organelles in reticulocytes, pages 151–180. Blood cell biochemistry. Plenum Press, 1990. 36

[292] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabsi. Hierarchical organization of modularity in metabolic networks. *Science*, 297: 1551–1555, 2002. 19

[293] R. J. Redfield. Is quorum sensing a side effect of diffusion sensing? *Trends Microbiol.*, 10:365–369, 2002. 34

[294] M. Ridley. *Evolution*. Blackwell Pub., 2004. 41

[295] P. Riggs. Expression and purification of maltose-binding protein fusions. *Curr. Protoc. in Mol. Biol.*, pages 16.6.1–16.6.14, 2001. 76

[296] R. J. Roberts, M. Belfort, T. Bestor, A. S. Bhagwat, T. A. Bickle, J. Bitinaite, R. M. Blumenthal, S. K. H. Detyarev, and D. T. F. Dryden. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.*, 31:1805–1812, 2003. 60, 61

[297] R. J. Roberts, T. Vince, J. Posfai, and D. Macelis. REBASE-restriction enzymes and DNA methyltransferases. *Nucleic Acids Res.*, 33:D230–D232, 2005. 60, 61, 62

[298] S. L. Rogers and V. I. Gelfand. Membrane trafficking, organelle transport, and the cytoskeleton. *Curr. Opin. Cell Biol.*, 12:57–62, 2000. 35

[299] D. Z. Rudner, Q. Pan, and R. M. Losick. Evidence that subcellular localization of a bacterial membrane protein is achieved by diffusion and capture. *Proc. Natl. Acad. Sci. U.S.A.*, 13:8701–8706, 2002. 29, 30

[300] M. E. Rumpho, F. P. D. adn J. R. Manhart, and J. Lee. *The structure and function of plastids*, chapter The kleptoplast, pages 451–473. Advances in photosynthesis and respiration. Springer, 2006. 36

[301] D. Sadava, H. C. Heller, D. M. Hillis, and M. Berenbaum. *Life: The Science of Biology.* W. H. Freeman, 2009. 37

[302] H. M. Salis, E. A. Mirsky, and C. A. Voigt. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.*, 27:946–950, 2009. 58, 59, 67, 69, 70, 73, 133, 134

[303] M. Sara and U. B. Sleytr. S-layer proteins. *J. Bacteriol.*, 182:859–868, 2000. 28, 29

[304] D. F. Savage, B. Afonso, A. H. Chen, and P. A. Silver. Spatially ordered dynamics of the bacterial carbon fixation machinery. *Science*, 327:1258–1261, 2010. 30

[305] G. Schatz and B. Dobberstein. Common principles of protein translocation across membranes. *Science*, 271:1519–1526, 1996. 37

[306] G. Schlosser and G. P. Wagner. *Modularity in development and evolution.* University of Chicago Press, Chicago, IL, 2004. 21

[307] H. N. Schulz and B. B. Jorgensen. Big bacteria. *Annu. Rev. Microbiol.*, 55: 105–137, 2001. 27

[308] B. Seyoum. *Export–import theory, practices, and procedures.* Taylor and Francis, 2008. 23

[309] J. A. Shapiro. Thinking about bacterial populations as multicellular organisms. *Annu. Rev. Microbiol.*, 52:81–104, 1998. 30, 48

[310] L. Shapiro, H. H. McAdams, and R. Losick. Why and how bacteria localize proteins. *Science*, 326:1225–1228, 2009. 27

[311] R. P. Shetty, D. Endy, and T. F. K. Jr. Engineering BioBrick vectors from BioBrick parts. *J. Biol. Eng.*, 2:5, 2008. 57, 59, 72, 75, 133

[312] J. M. Shively. *Complex intracellular structures in prokaryotes.* Microbiology monographs. Springer, 2006. 28

[313] J. Shoemaker-Galloway. Cargoshell collapsible shipping containers: a greener and flatter way to transport goods. [http://www.triplepundit.com/2010/02/cargoshell-collapsible-shipping-containers-a-greener-and-\
flatter-way-to-transport-goods/], 2010. 24

[314] N. Shubin, C. Tabin, and S. Carroll. Fossils, genes, and the evolution of animal limbs. *Nature*, 388:639–648, 1997. 39

[315] R. Singh, D. Paul, and R. K. Jain. Biofilms: implications in bioremediation. *Trends Microbiol.*, 14:389–397, 2006. 116

[316] S. Sleight. BBF RFC 26: In-Fusion BioBrick assembly. [http://dspace.mit.edu/handle/1721.1/46328], (Archived by WebCite at http://www.webcitation.org/5u3ZJyWmX), *DSPACE*, 2009. 57

[317] S. Sleight and B. A. Bartley. In-Fusion BioBrick assembly and re-engineering. *Nucleic Acids Res.*, 38:2624–2636, 2010. 57, 70

[318] U. B. Sleytr. Basic and applied S-layer research: an overview. *FEMS Microbiol. Rev.*, 20:5–12, 1997. 28

[319] U. B. Sleytr and T. J. Beveridge. Bacterial S-layers. *Trends Microbiol.*, 7: 253–260, 1999. 28, 29

[320] U. B. Sleytr, B. Schuster, and D. Pum. Nanotechnology and biomimetics with 2-D protein crystals. *IEEE Eng. Med. Biol.*, 22:140–150, 2003. 94

[321] M. H. D. Smit and J. V. Duin. Secondary structure of the ribosome binding site determines translational efficiency: A quantitative analysis. *Proc. Natl. Acad. Sci. U.S.A.*, 87:7668–7672, 1990. 141

[322] S. Srivastava. *Understanding bacteria.* Kluwer Academic Publishers, 2003. 31

[323] L. J. Stal. *Cyanobacterial mats and stromatolites.* Netherlands Institute of Ecology, 2000. 31

[324] C. Starr, C. Evers, and L. Starr. *Biology: today and tomorrow : with physiology.* CourseMate Series. Cengage Learning, 2009. 27, 29

[325] C. Steiner. *$20 per gallon: how the inevitable rise in the price of gasoline will change our lives for the better.* Grand Central Publishing. Grand Central Pub., 2009. 25, 26

[326] P. S. Steward, R. Murga, R. Srinivasan, and D. Beer. Biofilm structural heterogeneity visualized by three microscopic methods. *Wat. Res.*, 29:2006–2009, 1995. 116

[327] P. S. Stewart. Diffusion in biofilms. *J. Bacteriol.*, 185:1485–1491, 2003. 31

[328] P. S. Stewart and M. J. Franklin. Physiological heterogeneity in biofilms. *Nat. Rev. Microbiol.*, 6:199–210, 2008. 30, 33, 116, 117, 118, 130, 134

[329] T. Suzuki, T. Kanagawa, and Y. Kamagata. Identification of a gene essential for sheathed structure formation in *Sphaerotilus natans*, a filamentous sheathed bacterium. *Appl. Environ. Microbiol.*, 68:365–371, 2002. 31

[330] B. J. Swalla. Building divergent body plans with similar genetic pathways. *Heredity*, 97:235–243, 2006. 40

[331] J. Tang, D. W. Taylor, and K. A. Taylor. The three-dimensional structure of α-actinin obtained by cryoelectron microscopy suggests a model for Ca^{2+}-dependent actin binding. *J. Mol. Biol.*, 310:845–858, 2001. 93

[332] J. Tang, A. Ebner, B. Kraxberger, M. Leitner, A. Hykollari, C. Kepplinger, C. Grunwald, H. J. Gruber, R. Tampe, U. B. Sleytr, N. Ilk, and P. Hinterdorfer. Detection of metal binding sites on functional S-layer nanoarrays using single molecule force spectroscopy. *J. Struct. Biol.*, 168:217–222, 2009. 82, 143

[333] Y. Tashiro, M. Furubayashi, T. Morijiri, K. Suzuki, K. Yasuno, S. Matsuno, A. Katabami, K. Saito, and D. Umeno. *Escherichia coli* robots that freeze, smell, swell, and time-keep. *IET Synthetic Biology*, 1:41–43, 2007. 59

[334] K. A. Taylor and D. W. Taylor. Projection image of smooth muscle α-actinin from two-dimensional crystals formed on positively charged lipid layers. *J. Mol. Biol.*, 230:196–205, 1993. 91

[335] K. A. Taylor and D. W. Taylor. Structural studies of cytoskeletal protein arrays formed on lipid monolayers. *J. Struct. Biol.*, 128:75–81, 1999. 91

[336] L. M. Teixeira, A. Strickland, S. S. Mark, M. Bergkvist, Y. Sierra-Sastre, and C. A. Batt. Entropically driven self-assembly of *Lysinibacillus sphaericus* S-layer proteins analyzed under various environmental conditions. *Macromolecules*, 10:147–155, 2010. 81, 85, 143

[337] M. Terzer, M. Jovanovic, A. Choutko, O. Nikolayeva, A. Korn, D. Brockhoff, F. Zurcher, M. Friedmann, R. Schutz, E. Zitzler, J. Stelling, and S. Panke. Design of a biological half adder. *IET Synthetic Biology*, 1:53–58, 2007. 59

[338] F. Tian, G. Q. Chen, and Z. Wang. BBF RFC 46: Large-scale peptide modification on BioBrick proteins. [http://dspace.mit.edu/handle/1721.1/49521], (Archived by WebCite at http://www.webcitation.org/5u3ZBxsaK), *DSPACE*, 2009. 57

[339] R. D. L. P. Tomán. *A historical dictionary of the U.S. merchant marine and shipping industry: since the introduction of steam.* Greenwood Press, 1994. 23

[340] P. N. T. Unwin and R. Henderson. Molecular structure determination by electron microscopy of unstained crystalline specimens. *J. Mol. Biol.*, 94:425–440, 1975. 94, 109

[341] E. E. Uzgiris and R. D. Kornberg. Two-dimensional crystallization technique for imaging macromolecules, with application to antigen-antibody-complement complexes. *Nature*, 301:125–129, 1983. 91, 96

[342] T. Uzzell and C. Spolsky. Mitochondria and plastids as endosymbionts: a revival of special creation? The similarities between cellular organelles and prokaryotes are probably primitive features that arose independently from a common ancestor. *American Scientist*, 62:334–343, 1974. 36

[343] J. Valpuesta, J. Carrascosa, and R. Henderson. Analysis of electron microscope images and electron diffraction patterns of thin crystals of Ø29 connectors in ice. *J. Mol. Biol.*, 240:281–287, 1994. 105, 107

[344] C. D. von Dohen, S. Kohler, S. T. Alsop, and W. R. McManus. Mealybug β-proteobacterial endosymbionts contain γ-proteobacterial symbionts. *Nature*, 412:433–436, 2001. 36

[345] F. J. Vonk and M. K. Richardson. Developmental biology: serpent clocks tick faster. *Nature*, 454:282–283, 2008. 40

[346] W. Kühlbrandt and D. N. Wang and Y. Fujiyoshi. Atomic model of plant light-harvesting complex by electron crystallography. *Nature*, 367:614–621, 1994. 93, 109

[347] A. Wagner. *The Origins of Evolutionary Innovations: A Theory of Transformative Change in Living Systems.* Oxford University Press, 2011. 42

[348] G. P. Wagner, M. Pavlicev, and J. M. Cheverud. The road to modularity. *Nat. Rev. Genet.*, 8:921–931, 2007. 19

[349] P. Walker and E. Wood. *The Open Ocean.* Life in the sea. Facts on File, 2005. 27

[350] A. E. Walsby. Gas vesicles. *Microbiol. Mol. Biol. Rev.*, 58:94–144, 1994. 28

[351] P. Walter. The membranes of a eukaryotic cell. *Am. Zool.*, 29:501–510, 1989. 26

[352] H. H. Wang, F. J. Isaacs, P. A. Carr, Z. Z. Sun, G. Xu, C. Forest, and G. M. Church. Programming cells by multiplex genome engineering and accelerated evolution. *Nature*, 460:894–898, 2009. 45, 57, 58, 59, 70, 133

[353] J. Warren. *Evaporites: sediments, resources and hydrocarbons.* Springer, 2006. 34

[354] J. D. Watson, T. A. Baker, S. P. Bell, A. Gann, M. Levine, and R. Oosick. *Molecular Biology of the Gene.* Pearson/Benjamin Cummings, 2008. 141

[355] D. Weibel, R. Derda, J. Norville, and G. Whitesides. Using layers of paper to create synthetic microbial communities. A simple approach for investigation of cellular communication, fostering collaboration in basic research and bringing science to the classroom. [`http://www.keckfutures.org/conferences/synthetic-biology/grantees.html`], 2010. 5, 6

[356] R. Weiss and T. F. Knight. Engineered communications for microbial robotics. In *DNA Computing in* 6th International Meeting on DNA Based Computers *(DNA6) (Vol. 2054).* Springer, 2000. 119

[357] T. Wendt, D. Taylor, K. M. Trybus, and K. Taylor. Three-dimensional image reconstruction of dephosphorylated smooth muscle heavy meromyosin reveals asymmetry in the interaction between myosin heads and placement of subfragment 2. *Proc. Natl. Acad. Sci. U.S.A.*, 98:4361–4366, 2001. 93

[358] M. Wheelis. *Principles of modern microbiology.* Jones and Bartlett Publishers, 2007. 32

[359] S. Whitelam. Control of pathways and yields of protein crystallization through the interplay of nonspecific and specific attractions. *Phys. Rev. Lett.*, 105: 088102, 2010. 81

[360] W. Winn and E. Koneman. *Koneman's color atlas and textbook of diagnostic microbiology.* Lippincott Williams & Wilkins, 2006. 28

[361] R. S. Wotton. The ubiquity and many roles of exopolymers (EPS) in aquatic systems. *Scientia Marina*, 67:13–21, 2004. 31

[362] K. D. Xu, P. S. Stewart, F. Xia, C.-T. Huang, and G. A. McFeters. Spatial physiological heterogeneity in *Pseudomonas aeruginosa* biofilm is determined by oxygen availability. *Appl. Environ. Microbiol.*, 64:4035–4039, 1998. 33

[363] T. O. Yeates, C. A. Kerfeld, S. Heinhorst, G. C. Cannon, and J. M. Shively. Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nat. Rev. Micro.*, 6:681–691, 2009. 28, 141

[364] Y. Yokobayashi, R. Weiss, and F. H. Arnold. Directed evolution of a genetic circuit. *Proc. Natl. Acad. Sci. U.S.A.*, 99:16587–16591, 2002. 58

[365] Y. Zhu, M. A. Eiteman, R. Altman, and E. Altman. High glycolytic flux improves pyruvate production by a metabolically engineered *Escherichia coli* strain. *Appl. Environ. Microb.*, 74:6649–6655, 2008. 57